DESIGN AND IMPLEMENTATION OF A SYSTEM FOR COLOR RECOGNITION FOR VISUALLY IMPAIRED PEOPLE

BY

MOHAMMED TAWFIQ SAMARA

A Thesis Presented to the

DEANSHIP OF GRADUATE STUDIES

KING FAHD UNIVERSITY OF PETROLEUM & MINERALS

DHAHRAN, SAUDI ARABIA

In Partial Fulfillment of the

Requirements for the Degree of

MASTER OF SCIENCE

In

COMPUTER ENGINEERING

NOVEMBER 2017

KING FAHD UNIVERSITY OF PETROLEUM & MINERALS
DHAHRAN 31261, SAUDI ARABIA

DEANSHIP OF GRADUATE STUDIES

This thesis, written by **MOHAMMED TAWFIQ SAMARA** under the direction of his thesis adviser and approved by his thesis committee, has been presented to and accepted by the Dean of Graduate Studies, in partial fulfillment of the requirements for the degree of **MASTER OF SCIENCE IN COMPUTER ENGINEERING**.

Thesis Committee

Dr. Yahya Osais (Adviser)

Dr. Jihad Al-Sadah (Co-adviser)

Dr. Tarek Sheltami (Member)

Dr. Mohamed Deriche (Member)

Dr. Louai Al-Awami (Member)

Dr. Adel Fadhl Ahmed
Acting Chairman of Department

Dr. Salam A. Zummo
Dean of Graduate Studies

15/5/18

Date

عمادة الدراسات العليا
DEANSHIP OF GRADUATE STUDIES

In Dedication To

my parents

Mr. and Mrs. Tawfiq and Yousra

and

my sons

Tawfiq and Haroun

This humble work is a sign of my love to you

ACKNOWLEDGMENTS

All praise and glory to Allah for giving me health, courage and patience to complete my master degree.

For most, I want to express my love and gratitude to my mother Yousra Samara and my father Tawfiq Samara not only for being my parents but also for their continuous and valuable support and for all their prayers made for me.

I want to express my love for my sons, Tawfiq and Haroun, and all my family members for their support and help in overcoming all the challenges I have faced during my study.

Moreover, I would like to express my sincere gratitude to my advisor Dr. Yahya Osais and co-advisor Dr. Jihad Al-Sadah for their continues guidance and immense knowledge as well as for my committee members Dr. Mohamed Deriche, Dr. Tarel Sheltami and Dr. Louai Al-Awami for their valuable comments and motivation during my thesis study.

Special thanks to all COE staff and faculty, especially who taught me during my master journey for their immense knowledge and valuable guidance.

Also, I would like to express my sincere gratitude to KFUPM for giving me the opportunity to continue my master degree.

My deep appreciation goes to FABLAB Dhahran for providing the appropriate environment including tools and materials needed for packaging the prototypes required for the thesis experimental work.

All thanks to my KFUPM friends and colleagues for their encouragement and providing a pleasant atmosphere for me. Special thanks to my friends Khaled Muhaisen, Mohammed Qannan and Ahmed Shaheen for their valuable time spent on performing experiments for my thesis work.

Finally, to the place where I came from, I dedicate my achievement to my lovely Palestine especially Gaza Strip that faces all the pain and suffering from both the occupation and the unjust siege.

TABLE OF CONTENTS

ACKNOWLEDGEMENT **v**

LIST OF TABLES **x**

LIST OF FIGURES **xi**

ABSTRACT (ENGLISH) **xiv**

ABSTRACT (ARABIC) **xvi**

CHAPTER 1 INTRODUCTION **1**

1.1 Motivation 1

1.2 Needs for Assistive Technologies 4

1.3 Problem Statement 5

1.3.1 Objectives 9

1.3.2 Methodology 10

1.4 Contributions 11

1.5 Thesis Overview 12

CHAPTER 2 BACKGROUND **13**

2.1 Human Senses 13

2.2 Mechanics of Hearing 14

2.3 Sound Synthesis 15

2.4 Visual Substitution 16

2.5 Importance of Color 17

2.6 Color Models . . . 17
2.6.1 The RGB Color Model . . . 18
2.6.2 The CMY Color Model . . . 18
2.6.3 The HSV Color Model . . . 19
2.7 Cyber-Physical Systems Approach for Building Embedded Systems 19

CHAPTER 3 LITERATURE REVIEW 22
3.1 Generic Assistive Technologies . . . 22
3.2 Assistive Technologies for Color Perception . . . 27
3.3 Commercial Products . . . 30
3.4 Discussion . . . 31

CHAPTER 4 SYSTEM MODELING 35
4.1 The Initial System Modeling . . . 35
4.1.1 Color Sensing Module . . . 35
4.1.2 Color Sonification Module . . . 36
4.2 The Revised System Modeling . . . 39
4.2.1 Color Transformation Module . . . 40
4.2.2 Initial Approach for Audio Synthesis . . . 41
4.2.3 Revised Approach for Audio Synthesis . . . 43
4.2.4 Final Approach for Audio Synthesis . . . 44

CHAPTER 5 SYSTEM DESIGN 46
5.1 The Hardware . . . 46
5.1.1 The Initial Prototype . . . 48
5.1.2 The Final Prototype . . . 54
5.2 The Software . . . 57
5.2.1 The Initial System Model . . . 57
5.2.2 The Revised System Model . . . 62

CHAPTER 6 COLOR ENDORSEMENT WEB APPLICATION 65
6.1 Introduction . . . 65

6.2 Web Application Architecture 66
6.3 Modification to System Design 68

CHAPTER 7 SYSTEM EVALUATION **70**

7.1 Introducing the Device Experiment 70
7.1.1 Task 1 72
7.1.2 Task 2 73
7.1.3 Task 3 73
7.1.4 Task 4 74
7.1.5 Task 5 74
7.1.6 Discussion 75
7.2 Color Matching Experiment 76
7.2.1 Discussion 77
7.3 Color Shades Experiment 77
7.3.1 Color Grouping Task 77
7.3.2 Color Shades Ordering Task 78
7.3.3 Discussion 79
7.4 Color Recognition Experiment 80
7.4.1 Discussion 80
7.5 Results 81

CHAPTER 8 CONCLUSION AND FUTURE WORK **82**

REFERENCES **85**

VITAE **92**

LIST OF TABLES

3.1 Literature summary 33

5.1 Color Sensor Comparison 49

5.2 Processing Units Comparison 52

5.3 Raspberry Pi 3 Model B Specifications 53

5.4 Functions Descriptions 64

7.1 Experiment 1: Experimental Data 75

7.2 Experiment 2: Experimental Data 77

7.3 Task 3a: Experimental Data 78

7.4 Task 3b: Experimental Data 79

7.5 Experiment 4: Experimental Data 80

7.6 Comparison with literature 81

7.7 Overall Results 81

LIST OF FIGURES

1.1 Typical Snellen chart to estimate visual acuity [1] 2

1.2 Pipeline of the research 8

2.1 Primary colors of light (see colored version) 18

2.2 Primary colors of pigments (see colored version) 18

2.3 HSV Color Model (see colored version) 19

2.4 Cyber-Physical Systems 20

2.5 Cyber-Physical Approach 21

3.1 The Note-Taker for VI students [2] 24

3.2 Jungil's Developed System [3] 25

3.3 FingerSight device with two speaker-based vibrators [4] 26

3.4 Colorino [5] 30

3.5 Color Teller [6] 30

3.6 Talking Thermostat [7] 31

3.7 iBill [8] 31

3.8 Assistive Technology Classification 32

4.1 Initial system Modeling 36

4.2 Inconsistent acoustic representation examples 39

4.3 Revised system modeling 40

4.4 Example of the audio signal for the initial approach (time domain, instantaneous frequency) 42

4.5 Example of the chirp signal for the revised approach (time domain, instantaneous frequency for several values) 43
4.6 Example of the chirp signal for the final approach (time domain, instantaneous frequency for several values) 45

5.1 Typical features of high-end color indicator 47
5.2 Preliminary design for the pen 47
5.3 Color Sensors (one with TCS3472 driver and one with TCS230 driver) 49
5.4 Diffused RGB LED 50
5.5 Anker A7910011 Pocket Bluetooth Speaker 51
5.6 Processing Unit (Raspberry PI 3, Beaglebone Black, Arduino Mega) 52
5.7 Anker Powercore 10000mAh Power Bank 53
5.8 The system architecture 54
5.9 The initial Prototype 55
5.10 The revised system architecture 56
5.11 The final prototype 56
5.12 Initial Matlab Simulink Model 57
5.13 Building Flora color sensor interface 58
5.14 Finite state machine for RGB readings 59
5.15 Sound Generation and Playback FSM 59
5.16 Sound Preparation 60
5.17 Revised Matlab Simulink Model 60
5.18 LCD Subsystem 61
5.19 Building LCD interface 62
5.20 Functional Decomposition Diagram 63

6.1 Color Endorsement Web Application 66
6.2 Online Service Architecture 67
6.3 Final Matlab Simulink Model 68
6.4 ThingSpeak Simulink Block 69

7.1 Artifial fruits for the experiment 71
7.2 The VI subject performing experiment 1 72
7.3 Blindfolded subject performing experiment 1 73
7.4 Blindfolded subject performing experiment 2 76
7.5 Blindfolded subject performing color grouping task 78
7.6 Blindfolded subject performing shades ordering task 79

THESIS ABSTRACT

NAME: Mohammed Tawfiq Samara

TITLE OF STUDY: Design and Implementation of a System for Color Recognition for Visually Impaired People

MAJOR FIELD: Computer Engineering

DATE OF DEGREE: November 2017

With the advances made in computing in recent times, we are witnessing a growing interest in developing systems to help visually impaired people integrate within the society. Such systems are needed to enhance the quality of life for the visually impaired people. A major social challenge facing visually impaired is that of recognizing colors. The perception of colors is very important in interacting with people and the surrounding environment. This thesis presents the design and implementation of a real-time embedded system that can help visually impaired people recognize colors, interact, and take decisions based on their perception of colors. The hardware part of the system is a pen-like device, which can detect color and generate a language-independent auditory signal representing the HSV values of the recognized color. Further, a web application for color endorsement *is developed and deployed. This web application aims to include sighted people to help visually impaired people integrate into society by endorsing colors recognized by them using the proposed system. A set of experiments on the new system has been performed with the help of visually impaired and blindfolded participants. The outcomes of the experiments confirm the theoretical foundations on which the system is based. In addition, the experimental data shows that the proposed design is easy to use, cheap to build, and produces very satisfactory results.*

Keywords – Assistive Technologies, Visually Impaired, Sonification, Color Perception, Signal Processing

ملخص الرسالة

الاسم الكامل: محمد توفيق سماره

عنوان الرسالة: تصميم وتنفيذ نظام تعرف على الألوان للأشخاص المكفوفين وضعيفي البصر

التخصص: هندسة الحاسب الآلي

تاريخ الدرجة العلمية: نوفمبر 2017

مع التقدم المحرز في الحوسبة في الآونة الأخيرة، نشهد اهتماماً متزايداً في تطوير نظم تساعد الأشخاص المكفوفين على الاندماج في المجتمع. هناك حاجة ماسة إلى مثل هذه الأنظمة لتعزيز وتحسين نوعية الحياة للأشخاص المكفوفين. إحدى التحديات الاجتماعية الرئيسية التي تواجه الأشخاص المكفوفين هو التعرف على الألوان. إن إدراكهم للألوان مهم جداً في التفاعل مع الناس والبيئة المحيطة بهم. تقدم هذه الرسالة تصميم وتنفيذ نظام مدمج في الوقت الحقيقي يمكن أن يساعد الأشخاص المكفوفين على التعرف على الألوان والتفاعل معها واتخاذ القرارات استنادًا إلى إدراكهم للألوان. جزء العتاد في النظام هو جهاز يشبه القلم، والذي يمكنه اكتشاف اللون وإنشاء إشارة سمعية مستقلة غير لغوية ثلاثية القوام تمثل قيم HSV للون المكتشف. علاوة على ذلك، فقد تم تطوير ونشر تطبيق ويب لتأييد الألوان ومصادقتها. يهدف تطبيق الويب هذا إلى تضمين الأشخاص المبصرين لمساعدة الأشخاص المكفوفين على الاندماج في المجتمع من خلال اعتماد ومصادقة الألوان المكتشفة باستخدام النظام المقترح.

في الختام مجموعة من التجارب على النظام الجديد تم إجراؤها بمساعدة أشخاص مكفوفين وأشخاص مبصرين معصوبي العينين. نتائج التجارب تؤكد الأسس النظرية التي بني عليها النظام. بالإضافة إلى ذلك، تظهر البيانات التجريبية أن التصميم المقترح سهل الاستخدام، وتطبيقه غير مكلف، وينتج عن استخدامه نتائج مرضية للغاية.

CHAPTER 1

INTRODUCTION

In this chapter, we introduce our thesis by showing the motivation and the need for building systems to assist visually impaired people. Then, we formulate the problem we are trying to solve and state the objectives of our study. Thereafter, the Cyper-Physical approach we followed to build the system are described. At the end, the achieved contributions from our study are documented.

1.1 Motivation

Visual impairment is the condition where the ability of a person to see is reduced in a way that cannot be corrected by usual means such as standard correction glasses or contact lenses [9]. Vision loss is examined using the Snellen chart shown in Figure 1.1. This chart is commonly used for testing and measuring vision clarity of a person at a given distance. This measure is referred by the vision acuity [10]. The notation of the vision acuity can be represented by d_1/d_2, where d_1 is the distance at which the tested person can see an object and d_2 is the

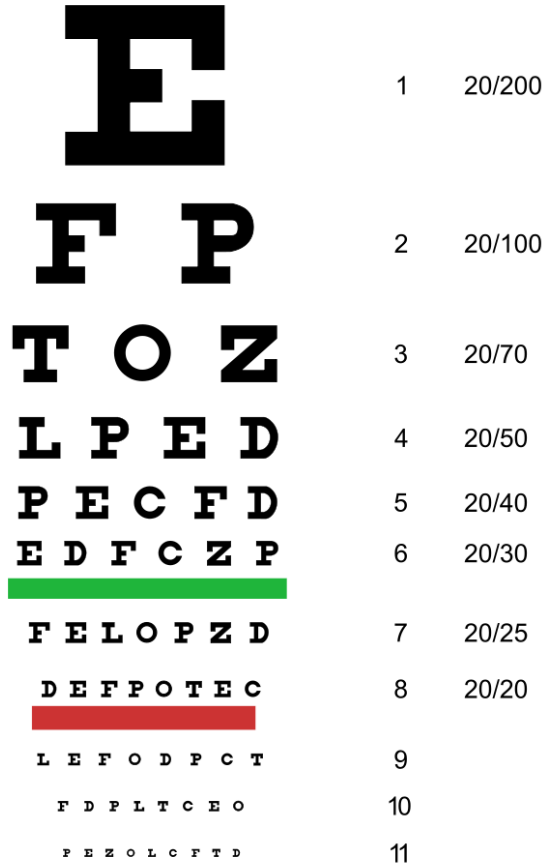

Figure 1.1: Typical Snellen chart to estimate visual acuity [1]

distance at which normal people can see that object. The distance is measured in feet or meters. In the united States, the normal visual acuity is described by 20/20 vision, which means the person can clearly see at 20 feet what should be normally seen at that distance. If the tested person's vision described by 20/70 vision, for example, it means that person needs to be as close as 20 feet to see what a normal person can see at 70 feet [11].

The World Health Organization (WHO) classifies visual impairments into the following categories:

- **Mild Vision Loss (Near-Normal Vision):** the vision acuity for the best eye ranges between 20/30 and 20/60.

- **Moderate Visual Impairment (Moderate Low Vision):** the vision acuity for the best eye ranges between 20/70 and 20/160.

- **Severe Visual Impairment (Severe Low Vision):** the vision acuity for the best eye ranges between 20/200 and 20/400.

- **Profound Visual Impairment (Profound Low Vision):** the vision acuity for the best eye ranges between 20/500 and 20/1,000.

- **Near-Total Visual Impairment (Near-Total Blindness):** the vision acuity for the best eye is more than 20/1,000.

- **Total Visual Impairment (Total Blindness):** the best eye does not have any perception of light. This category includes people who are born blind. This kind of visual impairment is referred to as *congenital blindness*.

In the United States, a person is considered *legally blind* when his vision cannot be corrected better than 20/200. Then, the person is considered for support programs provided by government or non-government organizations.

According to the WHO [12], there are 253 million Legally Blind and Visually Impaired (LBVI) people in the world. Further, 90% of them live in developing countries and about 65% of them are aged over 50 years old [13]. About 19 million children are visually impaired. According to the same statistics, 1.4 million children are blind for the rest of their lives and need personal care and vision rehabilitation.

From the above, it is very clear that there is an urgency for novel research on the needs of LBVI people and the impact of assistive technologies on their lives. This thesis represents an attempt in this direction.

1.2 Needs for Assistive Technologies

Loss of the visual sense can lead to a complete disability. Hence, immediate assistance is needed for people with vision loss. A Visually Impaired (VI) person also needs independence and wants to be integrated into the society. All these issues necessitate the need for sophisticated assistive devices which can help VI people overcome their disability.

Unfortunately, although there are a variety of commercial assistive technologies for VI people, they are still expensive and not widely available. In addition, most of them are not designed around the needs of VI people. Furthermore, manufacturers of such assistive devices target specific regions due to the standard of living and language barriers.

Activities such as personal hygiene, timekeeping, food preparation and consumption, household appliance usage, navigation, education, and shopping are top priorities for VI people. Color recognition and perception is typically not considered as an essential activity for VI people. However, color recognition can enhance the quality of lives of the VI people. It will help them learn about the colors of the essential objects around them like vegetables and fruits. It will also help them engage in discussions about interesting things in their lives such as the colors of their own cloths [14].

Based on our interviews with VI people, they lost interest in colors due to the lack of assistive devices for color recognition. In addition, they cannot discriminate between the different degrees of the same color. They also have a limited range of vocabulary whenever they are engaged in discussions about art and colorful objects.

This thesis is an attempt to tackle the hard problem of color perception for VI people. In addition, through preliminary experimental results, we are going to show the impact of our newly proposed technique on the lives of the VI people who use it.

1.3 Problem Statement

Conveying visual information using audio signals will make the world more accessible to VI people. For instance, they can appreciate art and recognize warning signals around them. They will also be more integrated into society and be able to speak new vocabulary, previously forgotten due to lack of usage.

One of the basic yet significant visual information is color. The physical reality of color is one dimensional: its frequency. Energy, wavelength and frequency are various representation of the same parameter. So, the visible color is the light perceived by the human eye with wavelength in the range of [390, 700] nanometers. Due to the physiology of the human eye, the color is perceived as three parameters due to existence of three pigment types in the eye. The relative response of these pigments is interpreted by the human brain as colors. In synthesizing colors, various combination of say red, green and blue lights mimics the same physiological equivalence to different colors. RGB model is only true in our eyes. Colors, from human perception, is represented using three-dimensional space RGB being a basic one. Another representation for the color is based on the human description for the color: HSV or Hue, Saturation and Value.

A common solution for conveying color information to VI people is by simply detecting the color, finding the closest named color based on the detected color values, and then pronouncing the name.

Such solution might help people who got visually impaired during their life because they have seen colors and interacted with them. However, congenitally blind people do not have any cognitive perception of colors because they have never experienced them. One common confusion is the mix between the color as a visual perception and the linguistic encoding. Orange the fruit has a unique color, but the name of the color is different. To the sighted persons they are the same. One common solution for conveying color information to VI people is by simply detecting the color, finding the closest named color based on the detected color values, and then pronouncing the name. Such solution might help people who got visually impaired during their life because they have seen colors and interacted with them. However, congenitally blind people do not have any cognitive per-

ception of colors because they have never experienced them. Hence, pronouncing the color names for them is nothing more than giving them an irrelevant piece of information that will be ignored or forgotten with time. Another defect in the linguistic encoding is the limitation of the language space compared to the color space. So, even if the VI person has knowledge of colors before losing the site, the colors names do not reflect the richness of the actual color spaces. This is a subtle point. People can differentiate thousands of shades of a color and color combinations. However, they can only name very few. Women are better in naming colors compared to men on average. So, giving the name of the color is a limitation and has no intrinsic value.

Another approach is communicating color information to the VI people through their other non-visual senses. It could be haptic feedback or non-speech sounds [1].

In this thesis, our goal is to design a human-centric device for assisting VI people to associate colors unique sound clips. The device has a special sub-system for converting color values to non-speech audio signals. This conversion process is referred to as *sonification*. We propose to represent each color using three audio signals which are played sequentially for the VI person. The advantage of our proposed approach is that it gives the VI person an acoustic signature which differs for each shade of color beyond color naming. This will help the VI people build a cognitive perception capability for colors.

[1] Non-speech sound is an audio feedback that does not use human speech

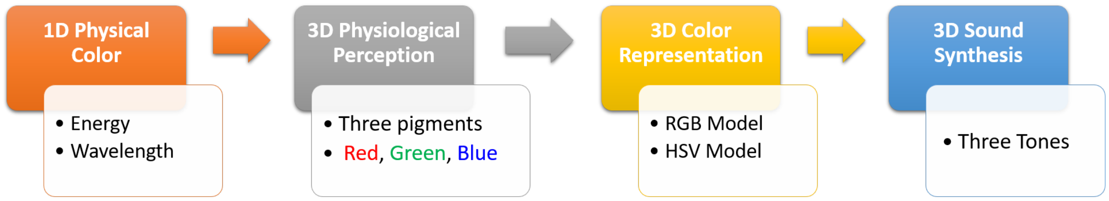

Figure 1.2: Pipeline of the research

Figure 1.2 shows the pipeline of our research, where we started by understanding the physical reality for colors in order to generate more intuitive and suitable audio signals for color representation.

It should be pointed out that giving the VI people the ability to like colors based on the acoustic effect of their representations might increase the gap with sighted people. The VI people might like the acoustic effect of a color that may not be visually comfortable for sighted people. To address this possible gap, we also propose a color endorsement service that will enable the VI people to share their desired colors with sighted people to get their ratings and feedback on the colors.

Most of the techniques sonify multiple low-level features at once which produce a large amount of information the VI person may not be able to comprehend. Color is a very rudimentary feature that deserves special attention and can be used to produce semantic distinction of objects.

Very valuable information can be communicated using colors only. For example, an orange can be described using its color only. This is why we only focus on sonifying color information in this thesis. Another main difference between our work and what has been done is representing a color by one acoustic signal rather than three acoustic signals. It is worth mentioning that evaluation of any sonification technique is challenging [15]. Thus, there is no sufficient data in the literature about how VI people can perceive colors using their other senses, even few of the proposed techniques use some form of sonification without experimental validation with real VI people.

1.3.1 Objectives

In this subsection we are introducing the main objectives for this thesis as follows:

1. Develop an assistive technology for the VI people to help them recognize colors using language-independent auditory signals as the color reporting mechanism. The auditory signals will vary according to the HSV[2] values of the detected color.

2. Design and implement a proof-of-concept prototype (pen) to be used for testing our proposed color reporting mechanism.

3. Assess the effectiveness of the proposed assistive technology through the implemented device by testing it with VI and blindfolded participants.

4. Develop an online service for color endorsement to help the VI people to be more integrated into society. Hence, they can know which colors are good for painting, furniture, and clothes.

[2]HSV is a color model and stands for the color attributes: Hue, Saturation and Value. More details are available on chapter 2.

1.3.2 Methodology

We are following the Cyber-Physical Systems (CPS) approach for building the embedding system of the assistive technology. The approach consists of three phases: modeling, design and analysis.

First, we start by modeling the problem's solution by developing a new color reporting mechanism to assist VI people in perceiving colors. This mechanism generates three consecutive audio signals for each detected color. The parameters of the generated audio signals are dependent on the HSV values of the detected color. A prototype is constructed to illustrate and evaluate the concept and mechanism, which is the design phase.

Secondly, we design a training procedure to introduce the device to VI people and help them get familiar with the auditory signals produced by the device. The VI people are asked to name common physical objects based on the audio signals they hear. Based on the analysis of the feedback we got from the VI people we revised our modeling and design stages of the problem.

Thirdly, we design and develop a web application for color endorsement. The main objective of this application is to help the VI people in making association between the acoustic representation of a color and what objects have that color. In which will help them become more independent.

Finally, the usability and effectiveness of the assistive technology are evaluated according to the feedback received during the training and testing stages.

1.4 Contributions

In the following, we describe our main contributions for this thesis:

- We have proposed and demonstrated a new technique for color sonification that helps VI people recognize colors. The new technique covers the 3D color space using three sequential sound signals while other works artificially assign a single tone to each color.

- Our technique has enabled the VI people to perceive (i.e. comprehend) and feel colors which encouraged them to socially interact with others using colors which means more integration into society. Thus, dealing with colors is a social need rather than a practical need for the VI people.

- Our technique represents the whole HSV color space while many similar studies collapse the color space into few colors.

- Compared to other similar studies, our technique needs less training time (24 minutes compared to 143 minutes) with higher success rate (99.18% compared to 76.77%).

- Our research has yielded new insights into the role colors can play in the life of VI people and nature of assistive technologies needed by them in order to be able to experience colors.

1.5 Thesis Overview

The rest of this thesis is organized as follows. The necessary background information is covered in chapter 2. Then, a detailed literature review is provided in chapter 3. After that, the system design is described in Chapter 4. The hardware and software details of the system implementation are mentioned in Chapter 5. Afterwards, the implementation details of the online color endorsement service are described in chapter 6. A set of performed experimental data are presented and discussed in Chapter 7. Finally, conclusions and directions for further research are given in Chapter 8.

CHAPTER 2

BACKGROUND

In this chapter, we present background information necessary to understand our research. First, the sensation and perception of humans are defined. Then, the mechanics of hearing and sound synthesis are explained. After that, visual substitution devices are introduced. Finally, an overview of the color models is given.

2.1 Human Senses

The humans have five sense organs: nose, skin, eyes, ears, and tongue. When one of these organs interact with the surrounding physical environment, the human body receives various stimuli that will be transformed into neural messages which will be transmitted to the brain [16]. Then the brain will use the received information and process it to interpret and understand the meaning of these messages. Therefore, it builds the experience of that stimulation. Although differentiating between sensation and perception is not always easy. Sensation is considered to be the initial contact between the sense organ and the environment. Interpreting and understanding the stimulation resulted from the contact is considered to be the perception.

The human has five senses: smell, touch, sight, hearing, and taste. In this thesis, we develop a technology that stimulates hearing sense so that the VI person can perceive colors.

2.2 Mechanics of Hearing

Hearing is possibly one of the most important human senses. People use hearing and sound to communicate information with each other and support their visual conversation. Sound is physically defined as a set of vibrations generated by a vibrant object [17]. These vibrations travel through a medium by causing a disturbance in the mediums molecules, which results in constructing mechanical waves. In our case, the medium is the air. Two main properties control the way the molecules move and how that movement sounds: (1) Frequency and (2) Amplitude.

Frequency is the movement rate of the molecules surrounding the vibrant object. It is measured by the number of cycles per second or hertz (Hz). The acoustic effect of the frequency in a sound is the pitch in that sound; the higher the frequency the higher the pitch. Amplitude is the maximum displacement in the surrounding molecules resulted by the vibrations of the object. The acoustic feel of the sounds amplitude is its intensity or loudness. The intensity of a sound is usually measured in decibel (dB).

It is worth mentioning that all sound waves travel at the same speed regardless of their frequency or amplitude. Higher frequency waves arrive more frequently to the human ear than lower frequency waves but both waves arrive at the same speed (1225 Km/hr at 20°C in sea level). That is because the lower frequency waves have longer wavelength than the higher frequency waves.

The human ear sensitivity is remarkable and it has wide range of dynamic hearing. It can detect one billionth of the atmospheric pressure, which is considered the low end of the dynamic range, which is called the threshold of hearing. On the other hand, the high end of the range is called the threshold of pain and is measured as 120 dB. This represents a dynamic range of hearing spanning 12 orders of magnitude across a frequency range from 20Hz to 20,000Hz.

2.3 Sound Synthesis

Sound is a set of vibrations traveling through a medium resulting mechanical waves. mechanical waves can be represented as continuous mathematical functions. The sum of trigonometric series can represent any continuous function, which is called Fourier Series. According to the concept of the Fourier Series, a sinusoid is considered to be the basic building block of a sound signal.

$$y = A\ sin(2\pi ft + \theta) \tag{2.1}$$

Where A is the signal's amplitude, f is the signal's frequency, and (θ) is the signal's phase. As we can see from equation (2.1), there are three parameters that can affect a sine wave. From an acoustic point of view, the amplitude represents loudness of a sound which is the user comfort level. While the phase of a sine wave does not affect its sound since it is a periodic function, different frequencies produce variety of sine wave sounds.

The Sound synthesis process consists of three main parts: inputs, process, and outputs. The outputs are synthesized audio that is generated by processing the inputs. If the synthesized audio is non-speech audio, then the process is called sonification.

2.4 Visual Substitution

Visual Substitution (VS) is about representing visual information such as color by using non-visual stimuli such as acoustic and tactile. Sensory Substitution Devices (SSDs) are the devices and aids that facilitate the VS application. There are several VS technologies that have been developed to help the VI people. Such technologies are Braille language, Text-to-Speech applications, normal or electrical long canes [18], navigation and guidance applications [19] [20] and many others. Wide variety of these SSDs are not widely used because there are expensive, troublesome and require a lot of training [21].

2.5 Importance of Color

Ulrich Beer, a psychologist, said about colors [22]:

> Seldom, surely, is the psychological part of an appearance in nature so great as it is in the case of color. No one can encounter it and stay neutral. We are immediately, instinctively, and emotionally moved. We have sympathy or antipathy, pleasure or disapproval within us as soon as we perceive colors.

Colors are considered to be descriptions for materials and they can be helpful in limiting the possible interpretations for objects. Also, colors can be used to provide visual cues to assist in navigation and organize traffic. Furthermore, colors can help people wear appropriate clothing by matching fabrics with visually comfortable colors. These are some of many examples that show the impact of using colors in our lives.

2.6 Color Models

Color models are standard representations of colors. In this section we are going to describe the most common color models: RGB, CMY and HSV models. Some of these color models are used for hardware and others are used for applications [23].

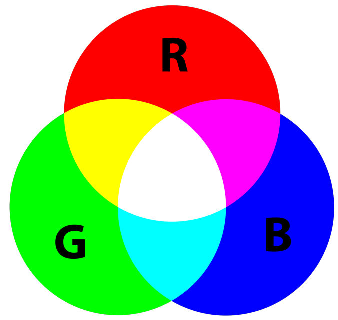

Figure 2.1: Primary colors of light (see colored version)

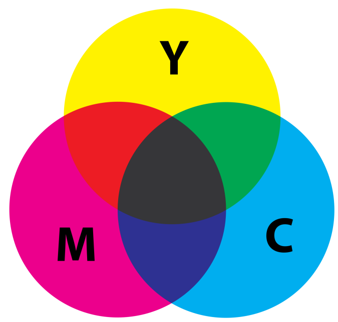

Figure 2.2: Primary colors of pigments (see colored version)

2.6.1 The RGB Color Model

The RGB model represents the colors as mixtures of three primary colors of light (Figure 2.1): Red, Green, and Blue. The white color can produced by mixing equal amounts of red, green and blue. The RGB color model is usually used with computer graphics and displays.

2.6.2 The CMY Color Model

The CMY model represents the color as mixtures of the primary colors of pigments (Figure 2.2): Cyan, Magenta, and Yellow. Mixing equal amounts of cyan, magenta, and yellow produces the black color. This color model is widely used in the printers' cartridges.

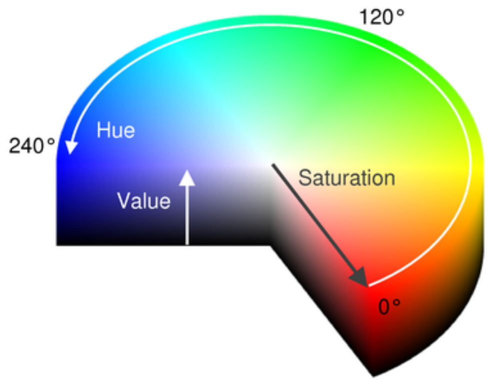

Figure 2.3: HSV Color Model (see colored version)

2.6.3 The HSV Color Model

We, as humans, describe the color by its hue, saturation and brightness. The hue, abbreviated by H, of a color is its attribute that describes a pure color, for example pure red, pure blue or pure yellow. While the saturation, abbreviated by S, of a color is the measure of how much the pure color is illuminated by the white light. Regarding the brightness, it is a subjective description and cannot be measured. Instead we are using the value, abbreviated by V, which represents the variation of the color from black to the color. The HSV color model cylinder is shown in Figure 2.3.

2.7 Cyber-Physical Systems Approach for Building Embedded Systems

A cyber-physical system (CPS) is a combination between a physical process and a computer system. The behavior of the system is defined by both its cyber and physical components. The CPS is an embedded system that monitors and

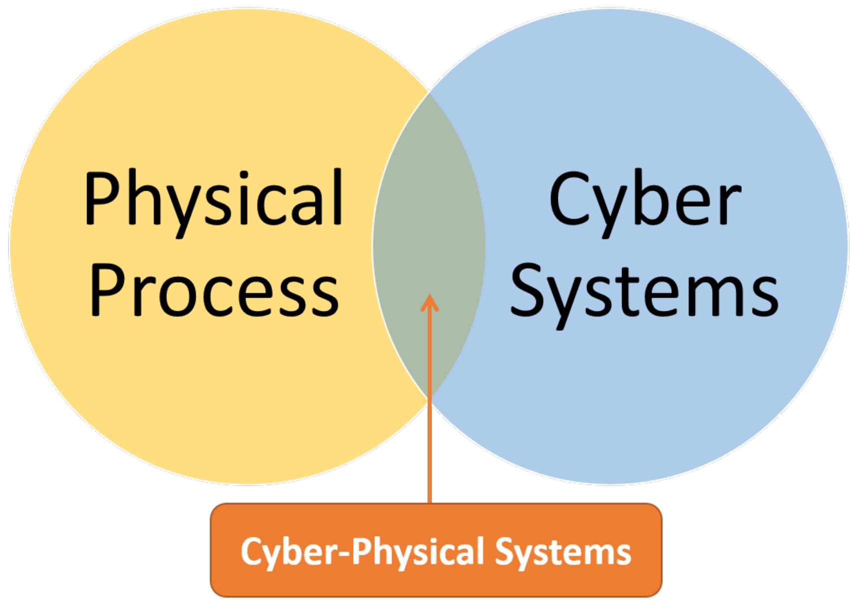

Figure 2.4: Cyber-Physical Systems

controls physical processes which affect the systems computations and hence its behavior. With CPS, it is very important to understand the interaction between the computational components and the physical components. It is about the intersection, not the union, between the physical processes and the cyber systems [24], as shown in Figure 2.4.

The process of building CPS consists of three major parts, modeling, design, and analysis. Modeling is to build mathematical representations for the physical process to gain deeper understanding for the system and its properties. Models specify what a system does. Design is the process of constructing the hardware and the software that implements the built mathematical models. It specifies how a system does what it does. Analysis is the phase where the CPS is tested and examined to see if the CPS meets its functioning requirements or fails to do so and know the reasons behind it. It specifies why a system does what it does or

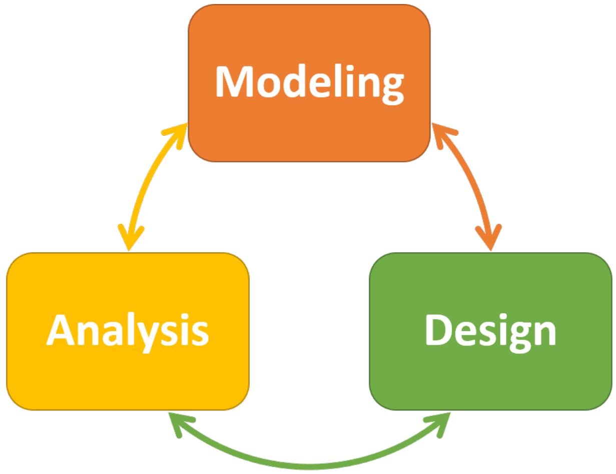

Figure 2.5: Cyber-Physical Approach

fails to do what it should do.

These three phases of the process must overlap and interact, as shown in Figure 2.5. The building process is an iterative process moves between the three phases. Usually, the building process starts with modeling to help understand the physical problem and develop solutions for that problem. The process may advance to the design phase by choosing suitable parts to build initial prototype to test and revise the models. During the analysis phase, the system models and designs are tested and evaluated.

CHAPTER 3

LITERATURE REVIEW

This chapter reviews the literature on the assistive technologies developed for VI people and gives a summary of the key milestones in the field. In the first section, general assistive technologies for VI people are viewed. After that, techniques for color perception are discussed and some commercial products are listed. At the end, a discussion that critique and compare between the technologies are provided.

3.1 Generic Assistive Technologies

Gracia et al. [25] proposed a system that uses commercial components such as Microsoft Kinect and Beaglebone to help visually impaired people in indoor navigation and obstacles avoidance. They used the Kinect to take live images and process them by the Beaglebone board to measure the distance from obstacles and then notifies the user via mechanical vibrations that get more intense as he/she gets closer to the obstacle.

Cazan et al. [26] proposed an application that can help the visually impaired by converting an image into rich sound signals that VI person can recognize. The application converts basic colors into musical tones. The first stage of the application is tested on sighted people and 45% of the sounds have been recognized.

Peng et al. [27] developed a system that can detect bar code and expiration date of a product using smart phones. The system first tries to detect the bar code and guide the user using non-visual cues, such as vibrations and sounds, for directions to undetected bar code. After that, the system searches for the expiry date of the product and apply an OCR algorithms to extract the date information and notify the user.

Taylor et al. [28] proposed guidance system in indoor environments using smart phones. The system acquires images using the smart phone's camera and sends them to a server in which a computer vision algorithm processes the images and try to locate the user and send him guidance through text-to-speech tools.

Balduzzi et al. [29] built a prototype that can detect faces and generate alerts for the visually impaired people if known faces are recognized. The user is notified using audio feedback by the name of the person if it exists in the dataset.

Lanigan et al. [30] proposed and developed a system that helps VI people have independent grocery shopping experience. Their system consists of a smart phone, Bluetooth headphones, bar code scanner and RFID tags. The barcode of a product is scanned and sent to the smartphone. It is checked for a product match. When there is a match, a product description is dictated to the VI person using a text-to-speech system.

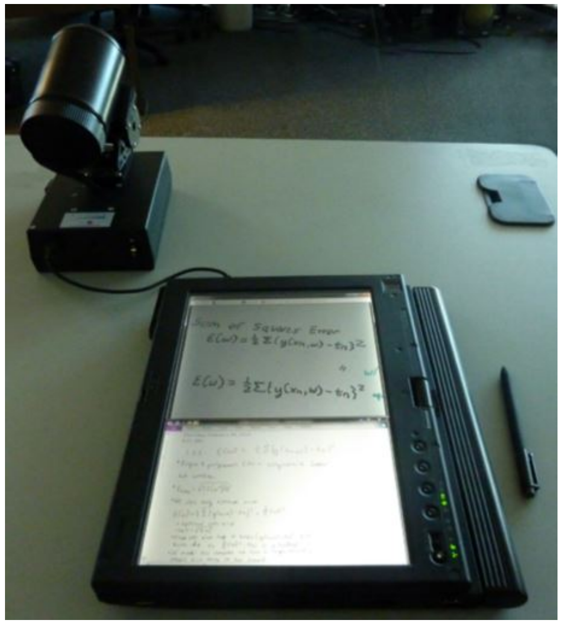

Figure 3.1: The Note-Taker for VI students [2]

Black and Hayden [2] implemented a note-taking device for VI students. The note-taker enables VI students to take notes in classrooms and keep up with the lecturers. The note taker consists of a tablet and video camera. The note-taker is shown in Figure 3.1. The note-taker runs a software on the tablet that helps the blind student see the board through the video camera feed and take notes at the same time. The student can control the video camera feed using software controls on the tablet, which help the student move around the board.

Jungil et al. [3] developed an assistive technology system for VI people to help them comprehend educational material. Their developed system contains software for instructors, software for low-vision students, and hardware for VI students. The instructor's software enables them to add and prepare educational material, which will be adjusted to suit low-vision and blind students. Then the adjusted material is wirelessly distributed to students. The low-vision students'

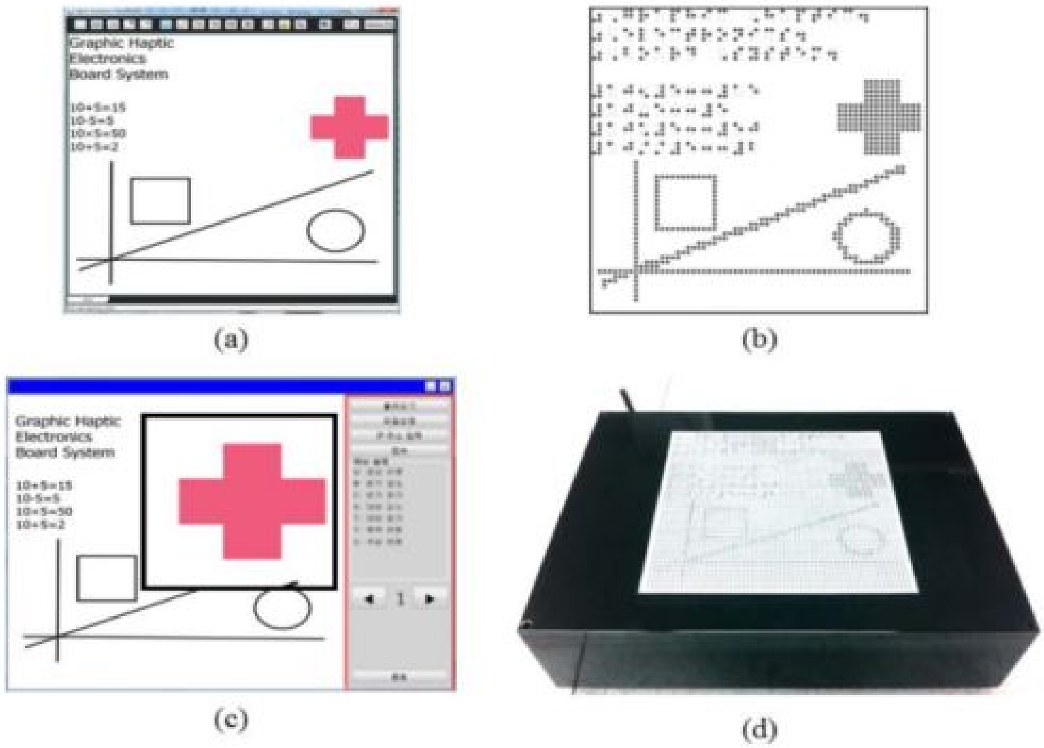

Figure 3.2: Jungil's Developed System [3]

software receives the material from the instructor's machine over a wireless network and adjusts it to the needs of the students. The VI students' hardware is a graphical haptic electronic board that consist of multi-layer tactile pins. The Braille representation of the education material is shown on board. Figure 3.2 shows Jungil's developed system.

Horvath et al. [4] developed an idea called FingerSight. It provides the finger with a haptic response according to the surrounding objects in a given environment. The idea is for the finger to be able to scan the environment and locate the objects based on their newly developed edge detection algorithm. They built several prototypes and performed proof-of-concept experiments on sighted blindfolded subjects. Their experiments showed that under active exploration, the

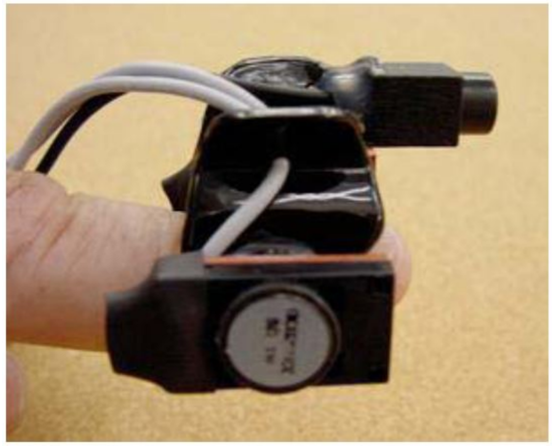

Figure 3.3: FingerSight device with two speaker-based vibrators [4]

FingerSight is capable of transmitting visual spatial information via haptic channels. Their final prototype is shown in Figure 3.3.

David Calder [31] addressed issues that concern the interfaces of developed assistive technologies for the VI people. The author emphasized that any assistive technology should not advertise for the user disability and it should not compromise any other senses such as hearing in order to provide information to the blind. Also, the assistive technology must be easy to learn otherwise it will mostly be rejected by the users according to Calder. In addition, he mentioned that the human brain can recognize certain sound-signature sequences in a similar way of recognizing musical tones. In this paper, he developed a basic prototype for detecting how far objects during navigation using sound substitution.

Pawluk et al. [32] considered issues that must be taken into consideration while designing assistive technologies for VI people. Such issues are diversity of population, tactile/haptic sensory substitution, participatory action design and real-world versus laboratory testing. Also, they emphasized that the test subjects must be categorized and the test should differentiate between early blind subjects who has no visual experience and between late blind subjects who has some visual experience.

3.2 Assistive Technologies for Color Perception

Abboud et al. [21] developed a novel visual to audio sensory substitution device for shapes and color information in an image. The system is called the EyeMusic. The EyeMusic scans the acquired image from left to right to create an auditory pixel-based waveform of the whole scene, and passes it to the subject using headphones. Each pixel in the image is assigned a musical note based on its vertical position. The timbre of that note represents the color value of the pixel, which is classified into one of the following six colors: white, blue, red, yellow, green and black. The musical instruments used are Choir for white, Brass instruments for blue, Reggae Organ for red, String instruments for yellow, and Rapman's Reed for green. The black is associated with silence. Their device was tested by twelve VI people and ten blindfolded people.

SonarX is a tool that convert color information in a digital image (pixel scanning) into musical octaves [33] [34]. While our research focuses on color information of an object, the tool also provides information about the location of objects in the image and their shapes. For the color information, the tool is based on the HSV color representation. The system sonifies the color using pitch, timbre and loudness for the hue, saturation, and value attributes, respectively. The tool was tested with eight VI subjects. Each test subject was presented by twenty trials made out of seven colors (red, orange, yellow, green, blue, purple, and violet).

Banf and Volker [22] [35] developed a modular sonification model to assist the VI people. While our research focuses only on color sonification, their model provides image pixels sonification. While our work focuses on the sonification of a single feature which is color, their work is about sonifiying low level features such as color, edges, orientation, and texture. For color sonification, the authors used musical instrument to represent colors based on the HSL color model. The instruments that are used for color sonification are choir for red, bagpipe for yellow, organ for green, strings for blue and flute for grayscale (i.e. white, black and gray).The saturation of a color is sonified by controlling the partial volumes of the playsing instruments while maintaining the relative volume ratio. Under a certain threshold of the saturation the color is considered gray and sonified using the flute. For the color lightness sonification, they change the pitch of the tone.

Bologona et al. [36] developed a system that transforms small part of a color image into sound. Their system is based on the HSL color model and is called Seeing Colors with an Orchestra (SeeColOr). The hue of a color is represented by instruments timbre. The instruments used are oboe for red, viola for orange, pizzicato violin for yellow, flute for green, trumpet for cyan, piano for blue, and saxophone for purple. The saturation is represented by the pitch of the selected instrument. When the luminance is dark, a double pass is mixed with sound resulting from the hue and saturation. Similarly, rather a singing voice is mixed with the resulting sound when the luminance is bright. The system is tested by ten blindfolded people using socks matching experiment based on their colors.

In 2013, a student [37] in art school invented coloring device for VI people. The device covered eight colors where the VI person can know the color using either braille labels on the colors or hearing the name of the color.

Cyborg Harbisson [5], a color-blind artist, came up with an idea to combine color with music (voice). His idea was implemented as the Eyeborg project which helped him detect colors by generating a tone to represent the color. He can distinguish between colors depending on how high or low the pitch of the generated sound.

Kees van den Doel [38] described how image colors can be translated unto an associated roughness encoded by varying scraping sounds. They represented the colors using one generated auditory signal for each image color. For a given color, the generated audio is a combination of several signal processing functions (Lowpass and Shepard filters) applied on a generated white noise. The filter parameters are dependent upon the hue, saturation and brightness of the image color.

Burch and Pawluk [39] developed a device that converts a 2-D, visually formatted diagram into its texture enriched haptic form. The key feature of the developed device is to transform color from a visually formatted diagram into different vibratory signals to mimic textures on a tactile graphic. The idea behind the device is to detect the light produced from the surface of a computer or any

Figure 3.4: Colorino [5]

Figure 3.5: Color Teller [6]

other illuminated screen. The detected light is then be converted into vibrations on the tip of a finger.

3.3 Commercial Products

In this section, we are listing some of the available products as assistive devices for LBVI people.

Colorino, shown in figure 3.4, is handheld color detecting device that can detect 150 color and report their names as a voice. It supports 20 different languages [5]. Again, another device that is language dependent whereas the presented solution in this research is not.

Color Teller is a compact device that help blind or color-blind people to identify colors [6]. The device operates using one button and can announce detected colors in three languages: English, French and Spanish. Figure 3.5 shows the Color Teller device.

Talking Thermostat is user friendly thermostat with audio playback [7]. It can be installed with multiple stage air conditioning systems. It plays information

Figure 3.6: Talking Thermostat [7]

Figure 3.7: iBill [8]

regarding day, time, indoor temperature, temperature setting and programming instructions. One of the talking thermostat is shown in figure 3.6.

iBill is talking banknote identifier device [8]. It has ultra-slim and compact design and it identify USD paper money quickly without depending on others. The iBill runs on a user replaceable AAA battery. iBill is shown in figure 3.7.

3.4 Discussion

According on the previous literature review, the implemented and published work of the assistive technologies can be classified into three main categories; Navigation, Shopping and Education, as shown in Figure 3.8. It is noticeable that how color recognition for VI people will add value to the available assistive technologies in these three categories. For example, in indoor navigation, rooms can be labeled by colors which will help VI people simply recognize indoor places depending on the colored labels.

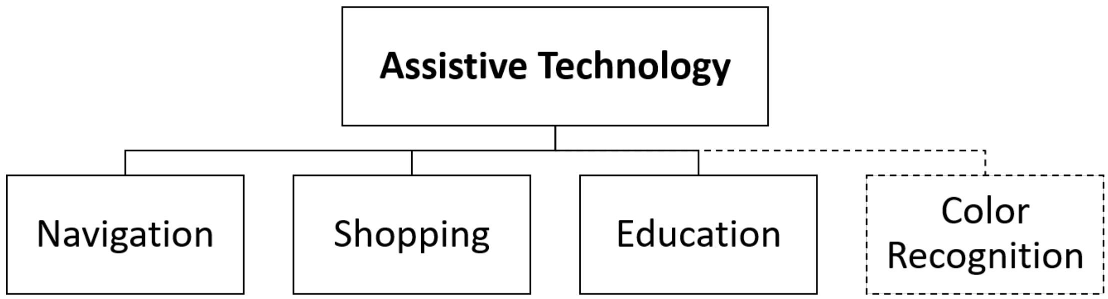

Figure 3.8: Assistive Technology Classification

Based on the key milestones of previously developed systems for color sonification, a system can be evaluated by its

- Color bandwidth,
- Training time,
- Response time, and
- Success rate.

The color bandwidth is the number of representable colors by the system. The training time is the time needed by test subjects to get familiar with the system and be ready to experiment with the system. The response time in an experiment is the time needed by VI subject to recognize a detected color since activating the system. The response time includes the color acquiring time, the processing time, and the sound synthesis. The success rate of a system is how many correct recognitions were given by test subjects using the system. For our research we are going to use these parameters to be the performance metric for our system.

Table 3.1: Literature summary

Year	**System Name**	**Color Space**	**Color Bandwidth**	**Training Time**	**Results**
2014 [21]	EyeMusic	HSL	5	120m-180m	81%
2013 [34]	SonarX	HSV	All	10m	53.13%
2012 [22]	-	HSL	All	240m-300m	96%
2008 [36]	SeeColOr	HSL	7	22.6m	94%

Abboud et al. [21] tested their device by twelve VI people and ten blindfolded people. The subjects were trained for two to three hours. For the color test, each participant was presented with fifteen color trials and then asked to give the name of each color. The success rate was 85.6% for the VI group and 75.7% for the blindfolded group. The overall success rate for both groups was 81%. The average response time for the blind group was 9.3 seconds and for the blindfolded group was 10.8 seconds.

Banf and Volker [22] [35] tested thier model on four VI subjects and the system's average success rate for the color test was 96%. The average response time was 2.1 seconds but the training time was four to five hours.

The SonarX tool [34] was tested with eight VI people. The subjects were trained for 10 minutes. Then, each test subject was presented by twenty color trials made out of seven colors (red, orange, yellow, green, blue, purple, and violet). The trials were classified correctly by percentage ranged from 29% for orange to 71% for violet. The overall average success rate was 50.14%.

SeeColOr [36] is tested by ten blindfolded people using socks matching experiment based on their colors. The participants were trained for around twenty three minutes. The average success rate for the system was 94%.

Table 3.1 summarizes the previous studies similar to our work and shows the results of their systems evaluation.

CHAPTER 4

SYSTEM MODELING

This chapter presents the details of our system modeling stage of the CPS approach for building the embedded system of the assistive technology. Further, several possible color sonification techniques are presented and investigated.

4.1 The Initial System Modeling

In order for the system to help the VI people recognize colors and perceive them using their hearing sense, we initially proposed a modeling of the system that consists of two modules as shown in Figure 4.1. The first module is a color sensing module while the other one is a color sonification module. Next, the details of the two modules are given.

4.1.1 Color Sensing Module

The puprpose of this module is to detect colors. The detected color is represented by an RGB color value. An RGB color value consists of three components: the

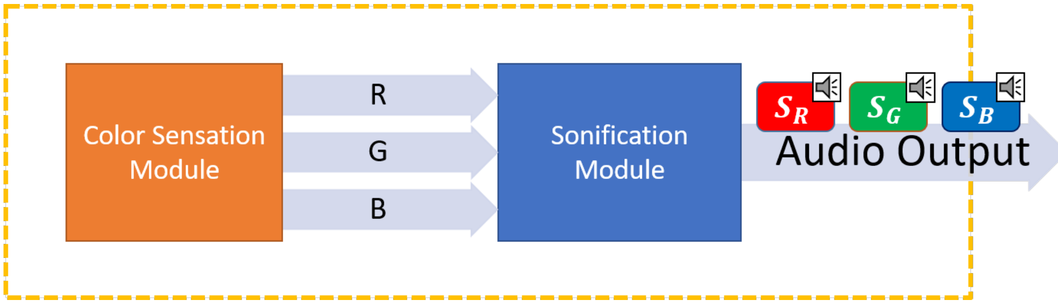

Figure 4.1: Initial system Modeling

red component, green component and blue component of the color. After the color is detected, its RGB value is passed to the color sonification module. RGB color space is used at this stage because all available color sensors report detected colors by their RGB value.

4.1.2 Color Sonification Module

The input of this module is the RGB color value generated from the color sensing module. Then, the module processes the passed RGB color value and generates three audio signals. Hence, it performs the sound synthesis process for the system. The synthesized sound consists of three sequential audio signals separated by one second of silence. The duration of each audio signal is also one second. These durations is selected based on our experimental observations over two sessions with two participants. The participants were introduced to several sound signals with the same frequency and different durations ranges between 100ms and one second. Both participants preferred the slower sounds (i.e. duration of one second).

For the synthesis process itself, two approaches based on the RGB color model are discussed next.

Color sonification is to represent the color using non-speech audio. Hence, we suggest to generate three sound signals for each detected color. Each sound signal represents a single color value. By changing the amplitude of a signal, we can generate sound signals with different levels of loudness. The loudness indicates how much red, green or blue is there in a given color. The following equations were used to generate three sounds representing a given color.

$$S_R(t; R) = R\ sin(2\pi f_r t) \tag{4.1}$$

Where S_R represents the audio signal of the red component. Similarly,

$$S_G(t; G) = G\ sin(2\pi f_g t) \tag{4.2}$$

$$S_B(t; B) = B\ sin(2\pi f_b t) \tag{4.3}$$

Where S_G and S_B represents the audio signals of the green and blue components, respectively. We propose two approaches for generating the red, green and blue audio signals. The first approach is to use the electromagnetic frequency values of the red, green and blue colors. On the other hand, the second approach utilizes the harmonics of the continuous Fourier series.

In the first approach, the electromagnetic frequency of pure red, green, and blue colors are used to generate three distinct sine waves (tones). The average physical frequency for each of the red (f_r), green (f_g), and blue (f_b) from the visible portion of the electromagnetic spectrum are computed and applied in equations

(4.1), (4.2) and (4.3), where $f_r = 442Hz$, $f_g = 566Hz$ and $f_b = 637Hz$.

In the second approach, the Fourier series harmonics are used. Because of the harmonics, three distinctive signals of the same fundamental frequency $f_\circ$ can be generated. The red, green, and blue audio signals are represented with square, triangle, and sawtooth waves respectively according to the following equations.

$$S_R(t;R) = \frac{4R}{\pi} \sum_{k=1}^{\infty} \frac{sin(2\pi(2k-1)f_\circ t)}{2k-1} \tag{4.4}$$

$$S_G(t;G) = \frac{8G}{\pi^2} \sum_{k=0}^{\infty} \frac{(-1)^k sin(2\pi(2k+1)f_\circ t)}{(2k+1)^2} \tag{4.5}$$

$$S_B(t;B) = \frac{B}{2} - \frac{B}{\pi} \sum_{k=1}^{\infty} \frac{(-1)^k sin(2\pi k f_\circ t)}{k} \tag{4.6}$$

After the initial implementation and testing of the two above approaches, two issues have caught our attention. The first issue was the use of the loudness (i.e. amplitude) of a signal as a way of differentiating between tones could generate unpleasant experience for VI people. High values of amplitude could generate high intensity audio signals. The second issue is that the acoustic representation of a color is not consistent among all possible colors. While most of the representations will have three tones, some of them will have one or two tones. For example, with pure red, blue, or green the subject will hear only one sound because the other two sounds will be eliminated because their RGB-dependent amplitude is equal to zero. In addition, the black color will only generate silence because the individual components of its RGB color value are all zero. Figure 4.2, shows three different audio representations that are inconsistent.

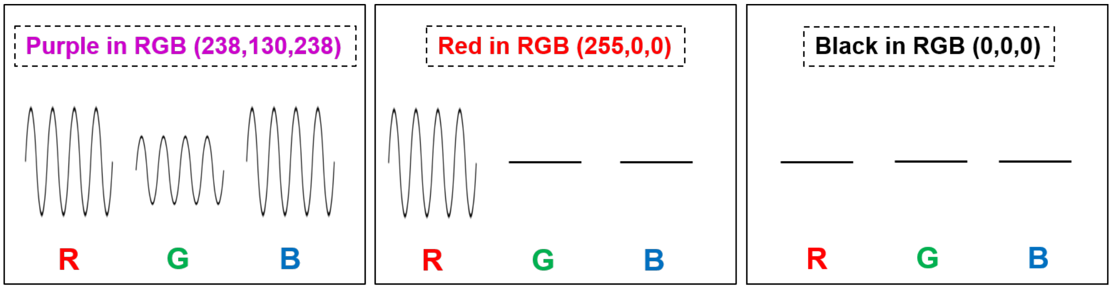

Figure 4.2: Inconsistent acoustic representation examples

Another issue is that the humans do not describe color using the RGB color model. Because of this issue, we needed to consider other color models as a base of the acoustic representations for colors. Due to the previous issues, we proposed to revise our modeling for the system which will be discussed in the next section.

4.2 The Revised System Modeling

Although the RGB color model matches how the human eye perceives colors, it is not how we differentiate between colors. While the RGB color model is best for hardware implementation either for producing colors, or analyzing them, we as humans describe color by its hue, saturation, and brightness. For the revised design, the HSV color model is used rather than the RGB color model because it is more intuitive and descriptive for humans [23]. For more details about the HSV color model please refer to Chapter 2. In order to use the HSV color model, one more module is added to the initial design. As shown in Figure 4.3, the added module is color transformation module. This module is used to convert the detected color from the RGB color space to the HSV color space.

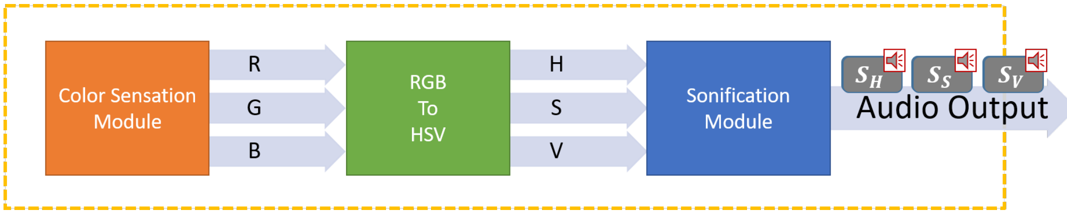

Figure 4.3: Revised system modeling

4.2.1 Color Transformation Module

After getting the RGB color value from the color sensing module, this module will convert the RGB values into the HSV values of the detected color. The conversion process is mathematically governed using equations (4.7) through (4.9) [23].

$$H = \begin{cases} \frac{60^\circ \quad |G-B|}{max(R,G,B)-min(R,G,B)}, & if \quad max(R,G,B) = R \\ \frac{60^\circ \quad |B-R|}{max(R,G,B)-min(R,G,B)} + 120^\circ, & if \quad max(R,G,B) = G \\ \frac{60^\circ \quad |R-G|}{max(R,G,B)-min(R,G,B)} + 240^\circ, & if \quad max(R,G,B) = B \end{cases} \tag{4.7}$$

$$S = \frac{max(R,G,B) - min(R,G,B)}{max(R,G,B)} \tag{4.8}$$

$$V = max(R,G,B) \tag{4.9}$$

Then, the HSV values are passed to the color sonification module. Color sonification approaches based on the HSV color model are discussed next.

Due to the acoustic representation inconsistency resulted from using the amplitude parameter to modulate the RGB color values, frequency parameter is used. Different frequencies produce wide variety of different tones. Therefore, the focus in the revised modeling is to develop a suitable and effective mapping between color values and the frequency of the auditory signals. Based on the study made by Cavaco and his colleagues [33], the frequency span must not be very narrow so that neighboring sounds would be easier to recognize. Thus, we should select the possible wider frequency range out of the human hearing range (i.e. 20Hz to 20,000Hz). Further, according to [40], the pleasant audible frequencies for the VI people should be less than 2500Hz although the range used in telephony is 300Hz to 3400Hz. Therefore, the upper frequency boundary (f_{max}) is selected to be 2200Hz. On the other hand, the lower boundary of the frequency range (f_{min}) is selected to be 200Hz even though human ear can hear 20Hz or higher frequencies and that is because the minimum frequency can be generated without harmonic distortion by general speakers is 200Hz [41]. Hence, we have selected the frequency range to be between 200Hz and 2200Hz.

4.2.2 Initial Approach for Audio Synthesis

At first, we started with a simple mapping of the HSV color values into three single frequency sinusoids, one sinusoid for each, of the HSV values. Using linear interpolation, the single frequencies are computed using (4.10), (4.11) and (4.12). Figure 4.4 shows an example of a tone generated using this approach. Smaller frequency values are used in this example and the following ones for the purpose of illustration.

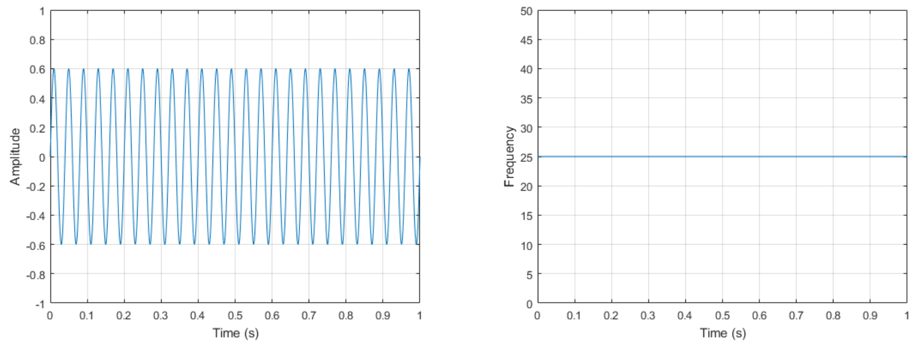

Figure 4.4: Example of the audio signal for the initial approach (time domain, instantaneous frequency)

$$f_h = (f_{max} - f_{min})\frac{h}{360°} + f_{min} \qquad where\ h \in [0°, 360°] \tag{4.10}$$

$$f_s = (f_{max} - f_{min})s + f_{min} \qquad where\ s \in [0, 1] \tag{4.11}$$

$$f_v = (f_{max} - f_{min})v + f_{min} \qquad where\ v \in [0, 1] \tag{4.12}$$

From these equations, we can see that we have unique mapping (i.e. one-to-one) between the 3D HSV color space and the 3D frequency space. Which means that each color shade has its own sound. Thus, our technique covers the whole HSV color space. However, This approach is a brute-force approach and results in uncomfortable audio signals irritating to the ear, especially after continuous listening to high-frequency tones. After numerous trials and errors, we improved this approach as in the following subsection.

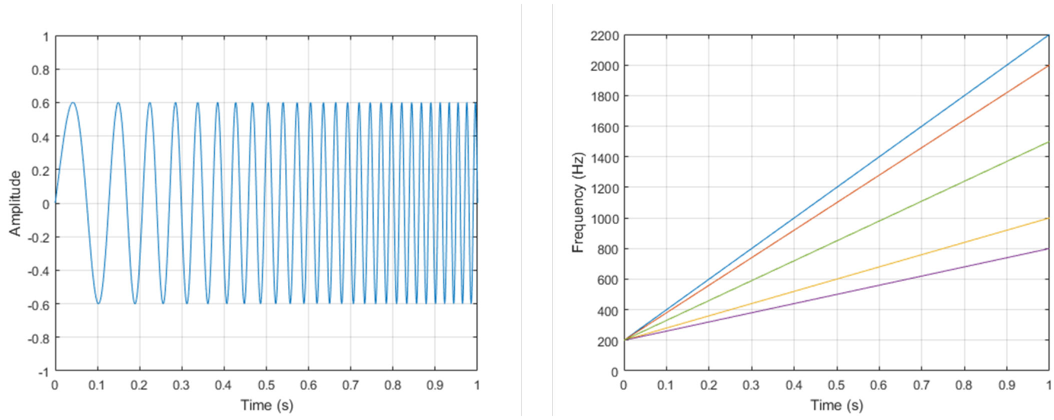

Figure 4.5: Example of the chirp signal for the revised approach (time domain, instantaneous frequency for several values)

4.2.3 Revised Approach for Audio Synthesis

In this approach, we also used (4.10), (4.11) and (4.12) to get the mapped frequencies but instead of generating sound of a single frequency, we generated a tone that rises with time. The tone starts from the lowest frequency boundary, which is 200Hz, and rises up to the calculated frequency, f_c. We implemented this idea by using a linear chirp signal, which is a signal whose frequency increases linearly over time, as shown in Figure 4.5. A tone is generated using equations (4.13) and (4.14).

$$k = \frac{f_c - 200}{t} \tag{4.13}$$

$$S = sin(2\pi t(200 + \frac{kt}{2})) \tag{4.14}$$

The initial testing for this approach showed that the generated tones are much more comfortable to the listener than the initial approach. However, this approach still presented some drawbacks. In particular, starting from the lower boundary makes the frequency change rate inconsistent especially when the calculated frequency is near the higher boundaries of the selected frequency range. This drawback make it harder to distinguish between colors lye at the end of the hue range. The next approach shows how it overcomes this drawback.

4.2.4 Final Approach for Audio Synthesis

In order to solve the inconsistency in the frequency change rate of the chirp signal, we are proposing to change the initial frequency of the chirping process. This time instead of starting from the lower boundary of our frequency range, we are proposing to start from the middle value of the frequency range, which is 1200Hz. In this case, the frequency will either increase or decrease over time until it reaches the calculated frequency, f_c. By doing this, we added more consistency to the change rate of the frequency. In addition, we achieved to add an acoustic effect to the generated tones that makes it even easier to the listener to differentiate between tones by giving away one bit of information through the rising or the falling acoustic effect of the chirp signal which means less mental effort needed to recognize colors. We used the following equations to generate the tones representing the detected color.

$$k = \frac{f_c - 1200}{t} \tag{4.15}$$

$$S = sin(2\pi t(1200 + \frac{kt}{2})) \tag{4.16}$$

Another modification is added to this approach which is performing the chirp-

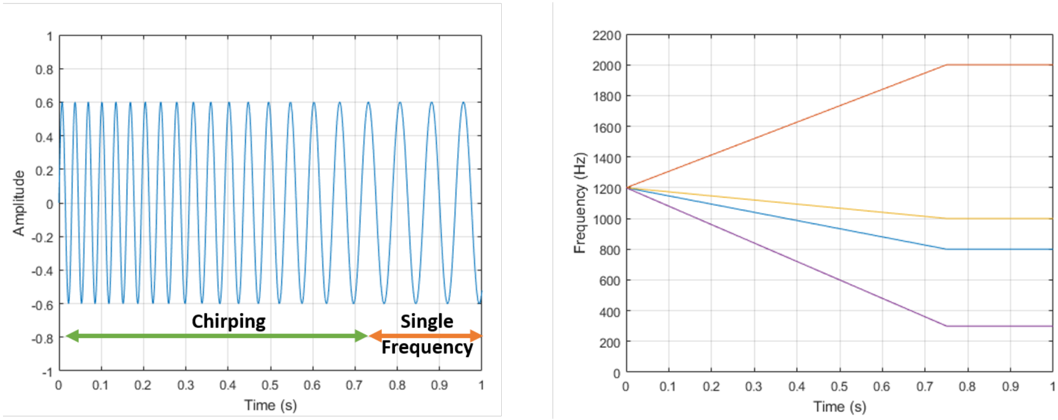

Figure 4.6: Example of the chirp signal for the final approach (time domain, instantaneous frequency for several values)

ing over three quarters of a second instead of the whole second. The remaining one quarter plays a single frequency tone of f_c. Thus, the final sound signal is represented by equation (4.17).

$$S = \begin{cases} sin(2\pi t(1200 + \frac{kt}{2})) & if \quad t \leq 3/4 \\ sin(2\pi f_c t) & if \quad t > 3/4 \end{cases} \tag{4.17}$$

This modification will help emphasize on the calculated frequency. Figure 4.6 shows an example of a tone generated using this approach.

Compared to the RGB approaches, our experience through the initial testing confirmed that the HSV color model is the best suited model for color sonification. The color representations are consistent because they are always generating three sequential tones. Also, the listener could easily recognize contrasted colors by only focusing on the first tone.

CHAPTER 5

SYSTEM DESIGN

This chapter represents the design stage of the CPS approach for building embedded systems. The system design consists of two parts: (1) Hardware Specification and Configuration and (2) Software Implementation which is model-based implementation.

5.1 The Hardware

As proof-of-concept, a pen incorporating our new mechanism (the final approach based on the HSV color model) is constructed. MATLAB Simulink is used in combination with a Raspberry Pi board to generate three consecutive audio signals for each detected color. Figure 5.1, shows the typical features [14] for a color indicator device which are considered in the design of the pen. A preliminary design for the pen is shown in figure 5.2.

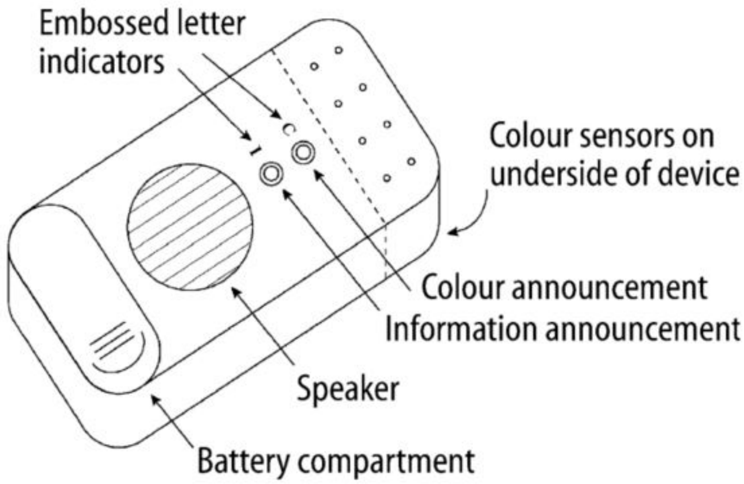

Figure 5.1: Typical features of high-end color indicator

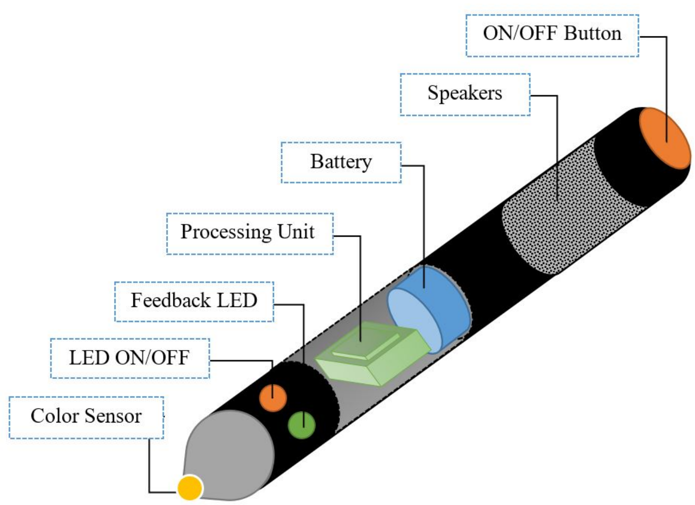

Figure 5.2: Preliminary design for the pen

5.1.1 The Initial Prototype

The device basically consists of the following parts:

- Color sensor for detecting colors,
- Push Button to initiates the process,
- Feedback LED to show the detected colors,
- Audio speaker to generate the auditory signals representing the detected color,
- Processing unit that have the capability to operate the required functionality, and
- Battery to power the device.

Color Detection Sensor

There were two color sensors available: Color sensor with TCS230 driver and the other one with the TCS3472 driver, shown in Figure 5.3. The TCS230 driver uses 8 by 8 array of photo-diodes and converts the color light into frequency represented by square wave. The frequency is directly proportional to light intensity. It communicates with any micro-controller using GPIO pins. On the other hand, the TCS34725 driver converts the color light into digital RGB values representing the sensed color. It is equipped with IR blocking filter for more accurate color detection. It is also provided with a white LED to illuminate on the color needed

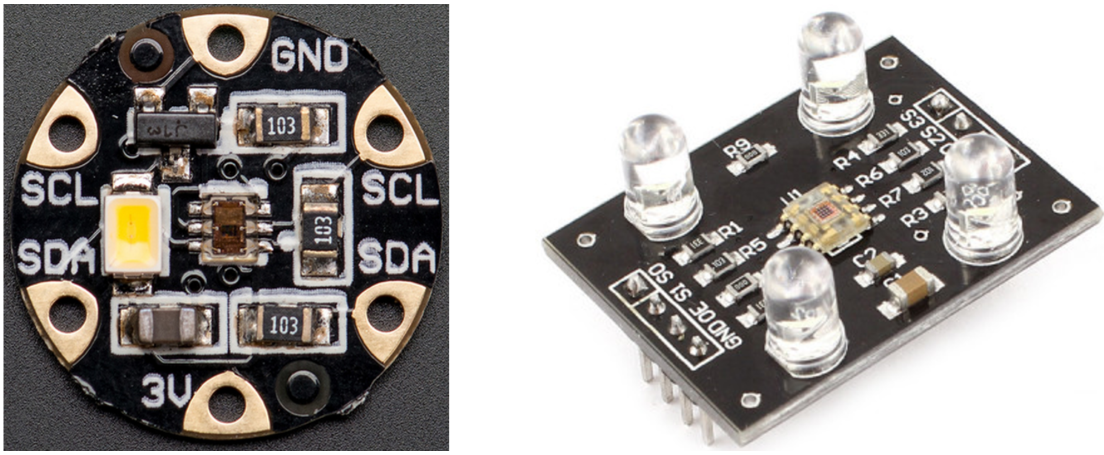

Figure 5.3: Color Sensors (one with TCS3472 driver and one with TCS230 driver)

Table 5.1: Color Sensor Comparison

Driver	TCS230	TCS3472
Methodology	Color light to frequency	Color light to digital converter
Filters	Red + Green + Blue	Red + Green + Blue + IR
Connection	GPIO	I2C
Price	\$6.5	\$8

to be sensed. It can be connected to any micro-controller board provided with I2C connectivity. Table 5.1 shows a comparison between the two sensors.

Based on the table, the color detection sensor we used is the Flora Color Sensor that uses the TCS3472 driver. And that is because it reports the RGB values of the detected color without any need of conversion and it has an IR blocking filter which make the sensor readings for the colors detection more accurate and closer to what humans can see of the colors. Also, the sensor is compact compared to the one using TCS230 driver.

The Flora Color Sensor can be connected to any micro-controller board provided with I2C connectivity pins. The connections is done as the following:

Figure 5.4: Diffused RGB LED

1. The 3V hole is connected to the 3V output pin.
2. The Ground hole is connected to the common ground.
3. The SCL hole is connected to the I2C Clock input pin.
4. The SDA hole is connected to the I2C Data input pin.

RGB Feedback LED

The feedback LED is needed to give an indication about the detected color. It will help the trainer or tester in making sure that the colors are correctly detected. For the feedback LED, we used 5mm RGB Diffused Common Cathode LED, shown in Figure 5.4. This LED have four pins; one pin for the common cathode and three pins for RGB colors. The common cathode is connected to the Ground. The RGB pins are connected to 3V output pins. The RGB pins are usually controlled using Pulse Width Modulation.

Figure 5.5: Anker A7910011 Pocket Bluetooth Speaker

Audio Speaker

The audio speaker is used to playback the generated tones representing the detected color. We used the Anker A7910011 Pocket Bluetooth Speaker, shown in Figure 5.5. We preferred this speaker because it is compact with high-quality playback sound and it has both bluetooth and 3.5mm audio jack connectivity. Also, it has rechargeable battery capable of 12 hours of playback time.

Processing Unit

We need this unit to handle the computations needed to convert the colors into non-speech audio signals. Also, It must have GPIO interface in order to be able to connect and use the color detection sensor and the RGB LED. It must have either a blutoohth connectivity or an 3.5mm audio jack for the Pocket Bluetooth Speaker to be used. There were three options: the Raspberry PI 3, the Beaglebone Black, or the Arduino Mega, shown in Figure 5.6. Table 5.2 shows a comparison

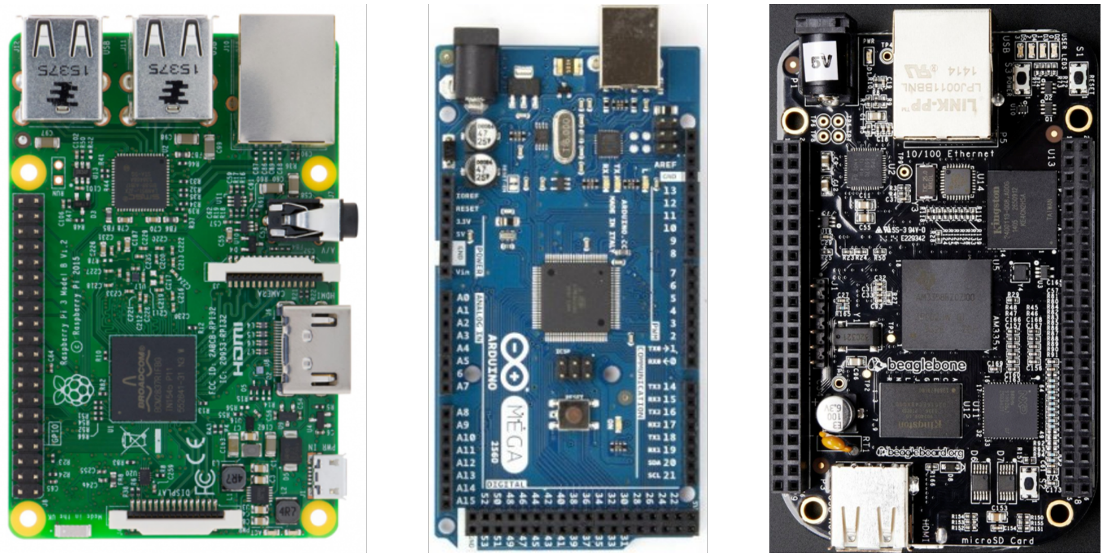

Figure 5.6: Processing Unit (Raspberry PI 3, Beaglebone Black, Arduino Mega)

Table 5.2: Processing Units Comparison

	Raspberry PI 3	**Beaglebone Black**	**Arduino Mega**
CPU Clock Rate	1.2 GHz Quad-Core	1 GHz	16 MHz
Flash Memory	1 GB	512 MB	256KB
Ethernet	Available	-	-
Wireless	802.11n + Bluetooth 4.0	-	-
GPIO	40 pins	92	70
Audio Jack	3.5mm Audio Jack	-	-
Price	\$35	\$45	\$38.5

between the three boards. Based on the table, the raspberry PI 3 is an easy winner because it is the most powerful one even though it is the cheapest. In addition, it has all the previous requirements and it is supported by Matlab Simulink. Table 5.3 shows detailed specification of the Raspberry PI 3 Model B board.

Table 5.3: Raspberry Pi 3 Model B Specifications

SoC	Broadcom BCM2837
CPU	1.2 GHZ Quad-Core ARM Cortex
Instruction Set	ARMv8-A
Memory	1 GB LPDDR2-900 SDRAM
Storage	micro-SD
GPU	Broadcom VideoCore IV @ 400 MHz
USB Ports	4 USB 2.0 ports
Ethernet	10/100 MBPS
Wireless	802.11n/Bluetooth 4.0
GPIO	40 pins
Audio	HDMI/3.5mm Audio Jack
Video	HDMI/Composite
Power Supply	micro-USB 5.1V (2.5A)

Figure 5.7: Anker Powercore 10000mAh Power Bank

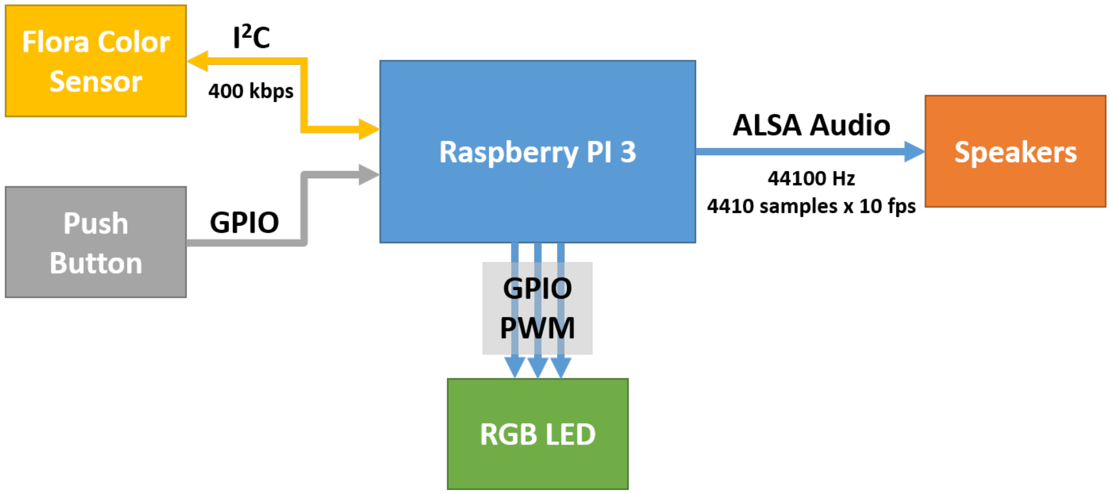

Figure 5.8: The system architecture

Battery

The battery should be able to steadily power up the selected processing unit while all needed peripherals are connected. For that, we selected Anker Powercore 10000 mAh power bank as the battery to power up our device. The power bank is shown in Figure 5.7.

The system architecture with all components connected is shown in Figure 5.8. An initial prototype of the device was constructed using the previous components. Then, it was tested during preliminary experiments with VI participants. The initial prototype is shown in Figure 5.9.

5.1.2 The Final Prototype

After testing the initial prototype, some issues were reported regarding its durability and usage. One of these issues is the RGB feedback LED, where its color mixture was not informative about the detected color. Another one is that the

Figure 5.9: The initial Prototype

connections might get loose if not handled carefully. Because of that, a new prototype is constructed.

While building the new prototype, we took into consideration all the issues. First, the RGB LED is replaced by an 128x160 TFT LCD, which gave a much better feedback for the detected color. The LCD is connected to the raspberry through the Serial Peripheral Interface bus (SPI). A revised system architecture incorporating the LCD is shown in Figure 5.10. Thereafter, The pen is 3D-printed to accommodate the color sensor and the push button. Finally, all the components are placed on a rectangular wooden base as shown in Figure 5.11. Therefore, the new prototype is more durable and easy to use and handle. Five prototypes were built to be used for the system evaluation.

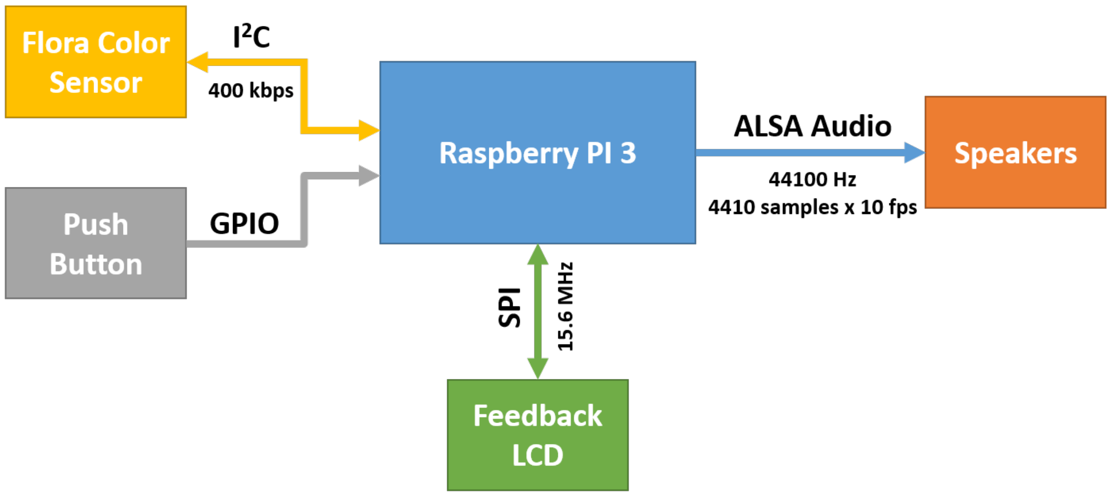

Figure 5.10: The revised system architecture

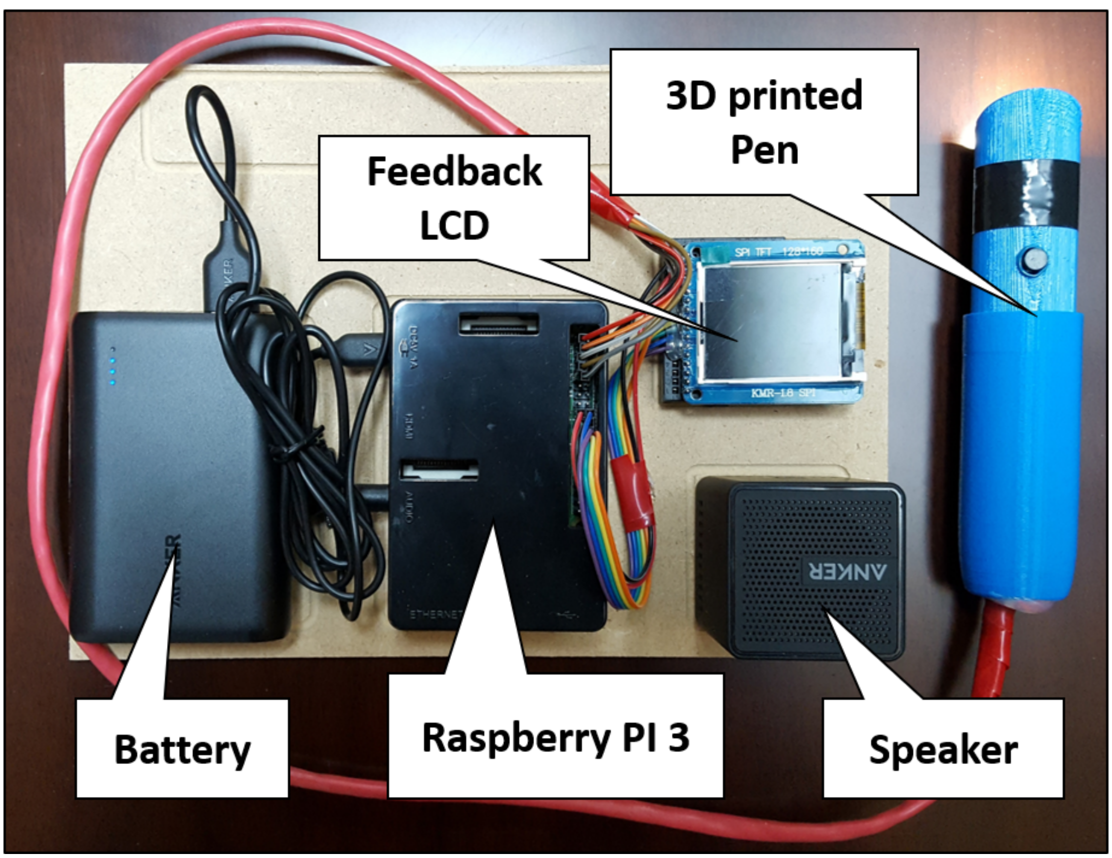

Figure 5.11: The final prototype

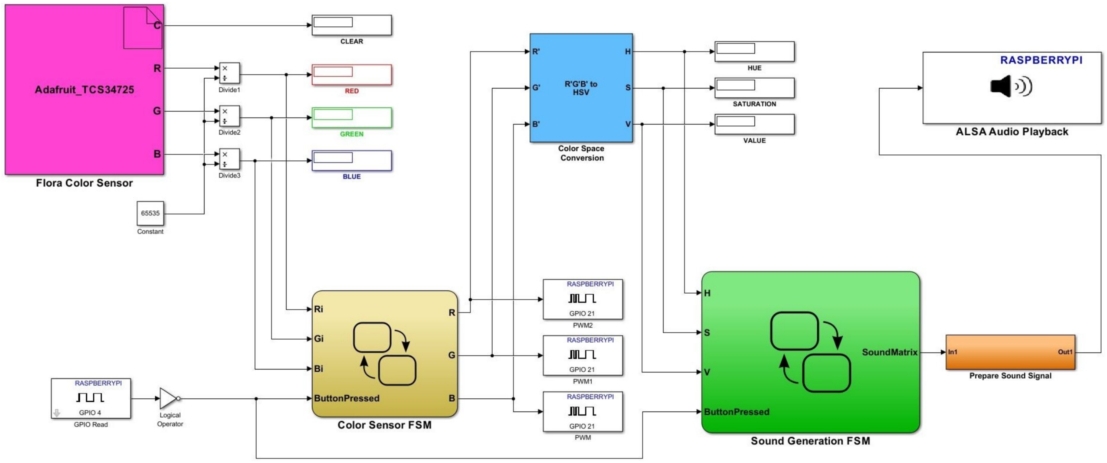

Figure 5.12: Initial Matlab Simulink Model

5.2 The Software

As for the software development, we are following model-based development approach to build the firmware of the device. Hence, we used MATLAB Simulink to build the system's model and install it on the Raspberry Pi 3 Model B. In the following subsection, we are going to discuss the different stages the software implementation had gone through.

5.2.1 The Initial System Model

The initial system model is the block model built for the initial prototype with the RGB feedback LED. The model is shown in Figure 5.12.

The pink block represents the flora color sensor. The block does not have any inputs and it has four outputs for clear, red, green and blue channels of the detected color. The clear channel is ignored. The sensor is connected to the raspberry through I2C pins, which is not supported in MTLAB Simulink 2016b

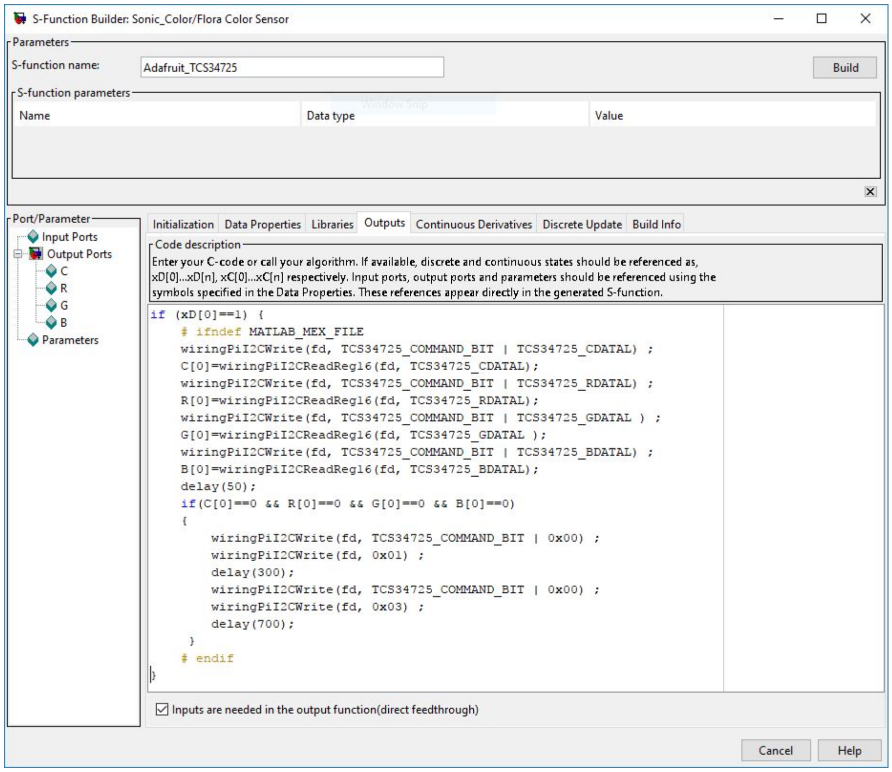

Figure 5.13: Building Flora color sensor interface

yet. We used the WiringPI library to develop and build an interface for the sensor using S-Function Builder, as shown in Figure 5.13. WiringPi is a PIN based GPIO access library written in C for the BCM2835 used in the Raspberry Pi. The sensor block provide RGB reading data every 100ms.

The yellow block regulates when to consider the RGB color readings. As shown in Figure 5.14, we used Finite State Machine (FSM) that passes the RGB reading data when the button is pressed. This FSM includes the Start, Wait, and Pass states. When the system starts the FSM moves from the Start state into the Wait state and when the button is pressed the latest RGB readings will be passed as output of this block. After six seconds the FSM will wait again for a button press.

The outputs of the FSM block is used to regenerate the detected color using

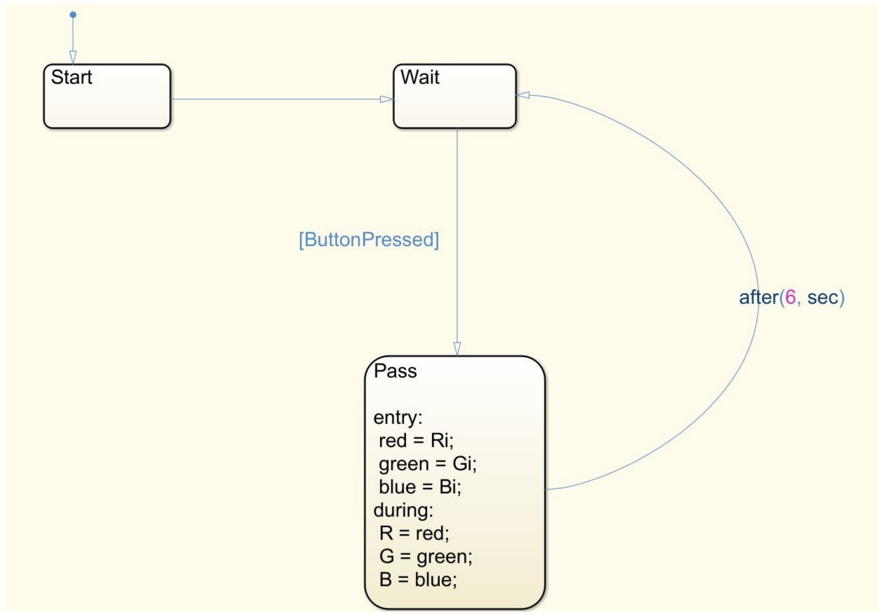

Figure 5.14: Finite state machine for RGB readings

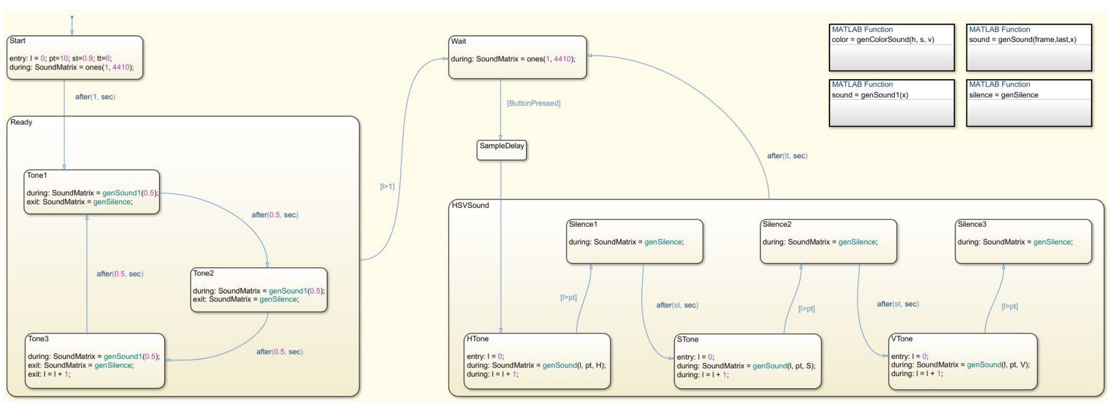

Figure 5.15: Sound Generation and Playback FSM

the RGB feedback LED. We used the Pulse Width Modulation (PWM) Simulink blocks to control the RGB LED pins. The PWM is a technique that enables digital pins to simulate analog signals using digital signals. The digital control creates a squared signal, switched on and off. When the signal is on, it simulates HIGH voltage and when it is off, it simulates LOW voltage. By changing the duty cycle of the squared signal we get analog voltage between LOW and HIGH voltages.

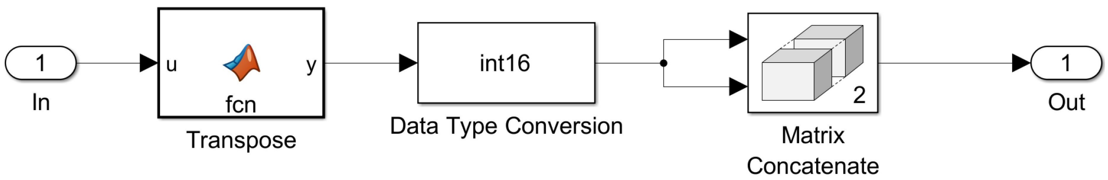

Figure 5.16: Sound Preparation

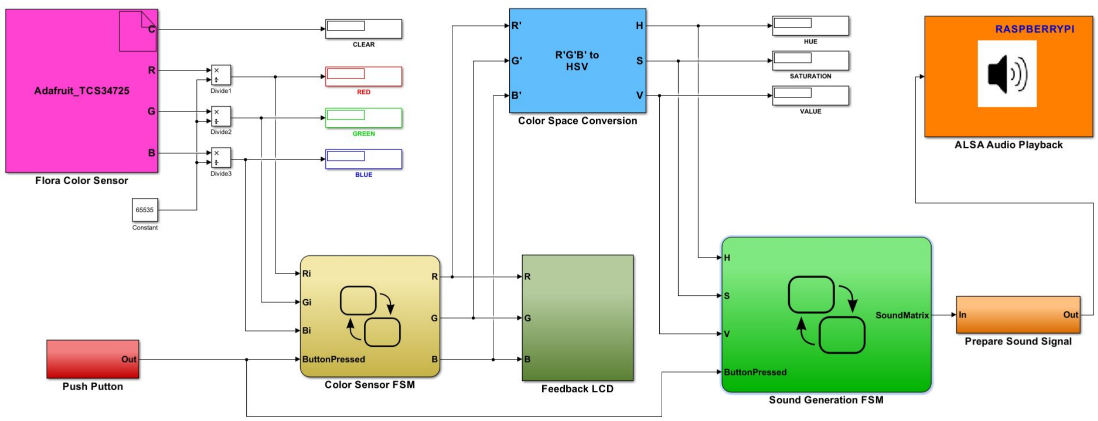

Figure 5.17: Revised Matlab Simulink Model

The selected RGB color readings are passed to the blue block as well. The blue block is a Simulink block that converts color readings from the RGB color space to the HSV color space. The green block encapsulates the FSM that generates and control the playback of the sensed color sound to the speakers. As shown in Figure 5.15, the FSM starts when the system model starts simulation. The FSM includes the Start, Ready, Wait, and HSVSound states. In the Start state, necessary initializations are performed and after one second it will move to the Ready state. The Ready state generate three audio signals to alert the user when that the device is ready and operable. After that, the FSM will wait for the user to press the button which will trigger different system events. When the button is pressed, the HSVSound state will process the HSV color values and

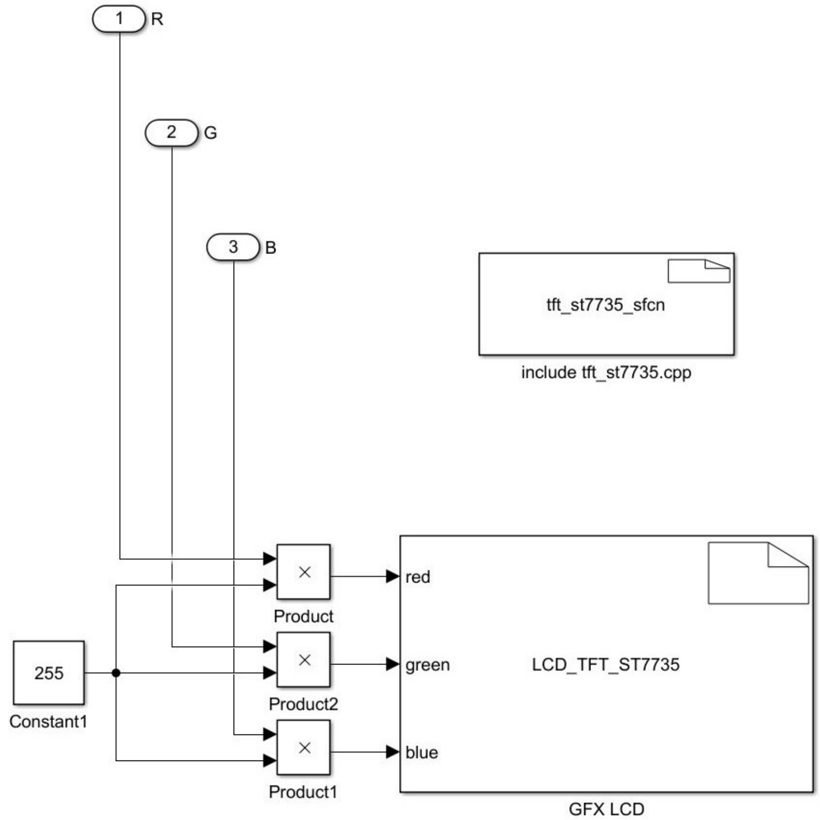

Figure 5.18: LCD Subsystem

produce three one-second tones separated by one second of silence. The generated sound samples are then prepared to be compatible with ALSA Audio Playback mechanism, as shown in Figure 5.16. The ALSA Audio card should receive two one-column matrices, one for the right channel of the speaker and the other for the left channel. The transpose function convert the one-row sound matrix into one-column matrix. Then, the resulted matrix is duplicated and concatenated to accommodate the ALSA Audio card requirements.

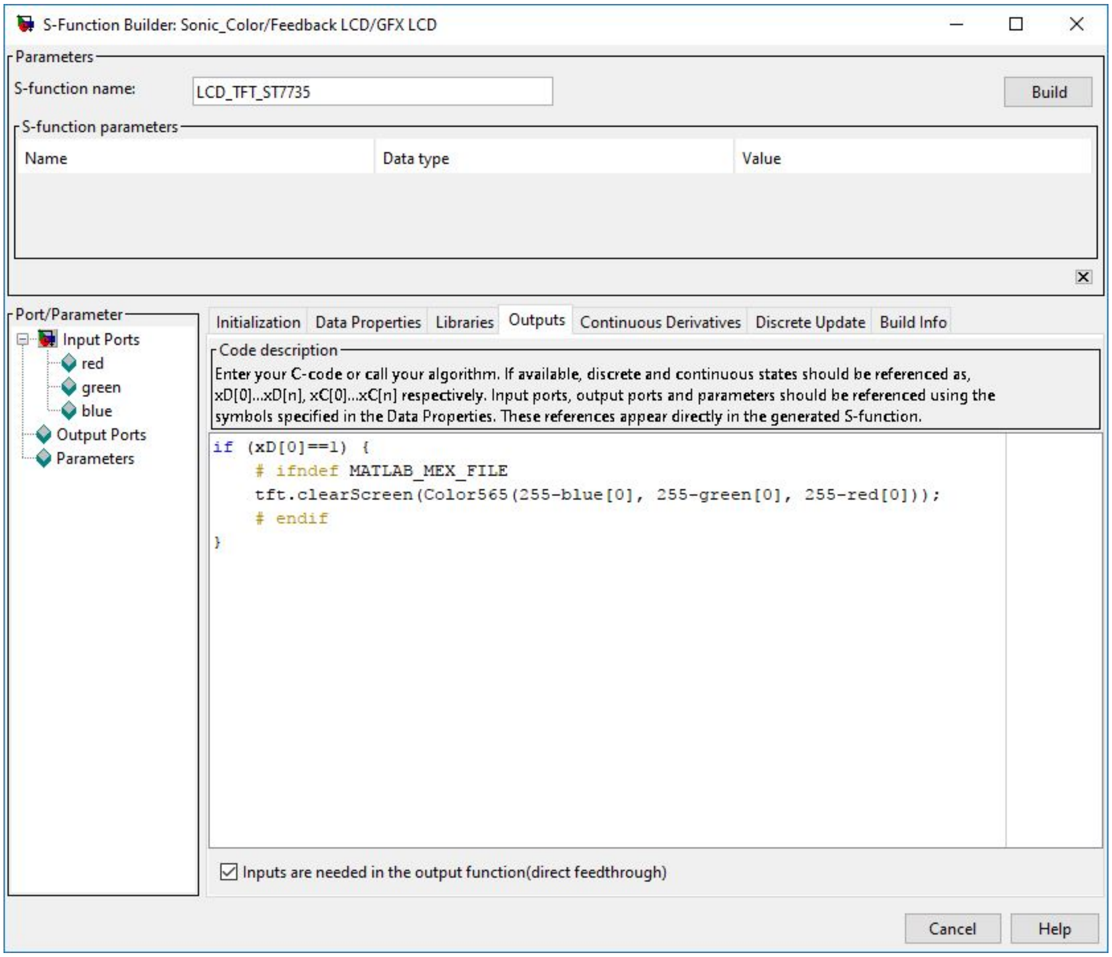

Figure 5.19: Building LCD interface

5.2.2 The Revised System Model

The revised system model, shown in Figure 5.17, incorporates the feedback LCD which replaced the RGB LED in the final prototype. The dark green block is an S-Function Builder block. Similar to the color sensor, We used WiringPI library to interface the LCD with the raspberry using SPI pins as shown in Figure 5.19. The LCD pixels controller is based on the CMY color model. To transform the color values from the RGB color space to the CMY color space, each component of the RGB color value is subtracted from 255 value to get the CMY color value.

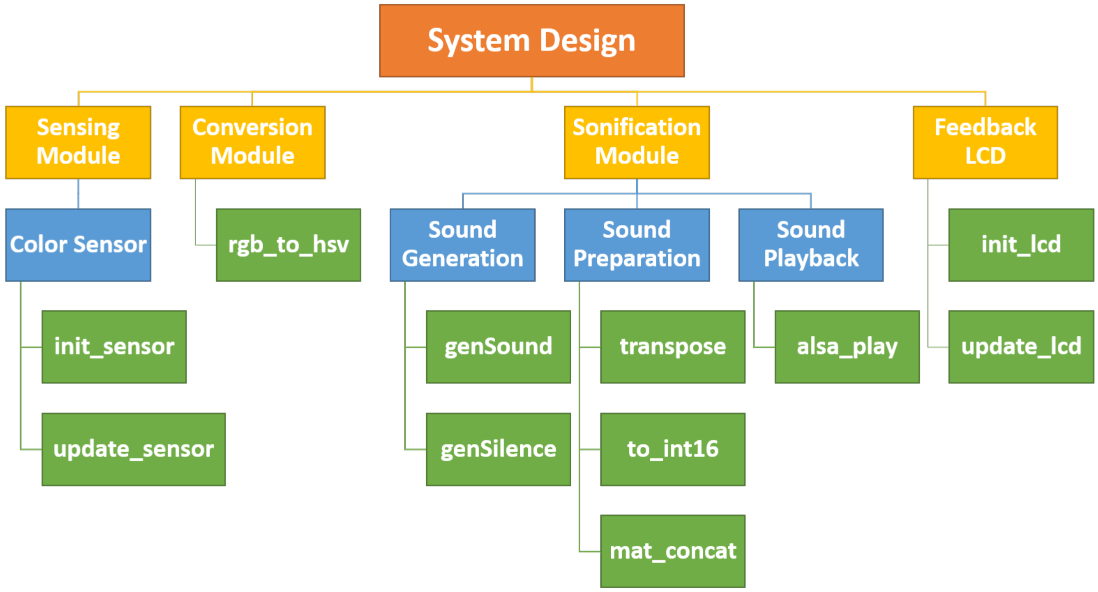

Figure 5.20: Functional Decomposition Diagram

Figure 5.20 shows the functional decomposition diagram of the final design of the system. As shown in the diagram, the system has three levels of abstraction. The first level includes the main components of the system (i.e. color sensation, conversion, and sonification) and the feedback LCD. The second level consists of the sub-systems of the first level components. For example, the sonification module contains three sub-systems: the Sound Generation sub-system, the Sound Preparation sub-system, and the Sound Playback sub-system. The third level includes the low-level functions of the sub-systems. Table 5.4 lists the low-level functions and their descriptions.

Table 5.4: Functions Descriptions

Function	**Description**
init_sensor	it initialize the sensor upon the device power on
update_sensor	it updates the sensor registers every 100ms
rgb_to_hsv	it converts the RGB color values provided by the sensors into the HSV color space
init_lcd	it initialize the feedback LCD upon the device power on
update_lcd	it updates the feedback LCD to reproduce the detected color every 100ms
genSound	it generates 1D sound matrix of size 4410 every 100ms based on the HSV color values
genSilence	it generates 1D silence matrix of size 4410 every 100ms
transpose	it converts the 1D sound matrix from 1 row to 1 column
to_int16	it converts the data type of the matrix into int16 as preparation for sound playback
mat_concat	it converts the 1D matrix into 2D matrix as preparation for sound playback
alsa_play	it plays the generated sound matrix through the ALSA sound driver of the raspberry pi

CHAPTER 6

COLOR ENDORSEMENT WEB APPLICATION

In this chapter, the web application design and development are discussed in detail. In addition, the required modifications to the system Implementation to accommodate the web application requirements are presented.

6.1 Introduction

The main objective of our system is to help the VI people in their daily life activities and to integrate them with society. Sighted people endorse colors based on their visual experience and preference. Using our system, we are trying to give VI people the ability to sense and perceive colors using their hearing sense. Therefore, we are expecting them to build a cognitive perception for colors using their acoustic representations. Hence, they will start endorsing colors based on that experience. They might like the acoustic effect of a color that is not visually

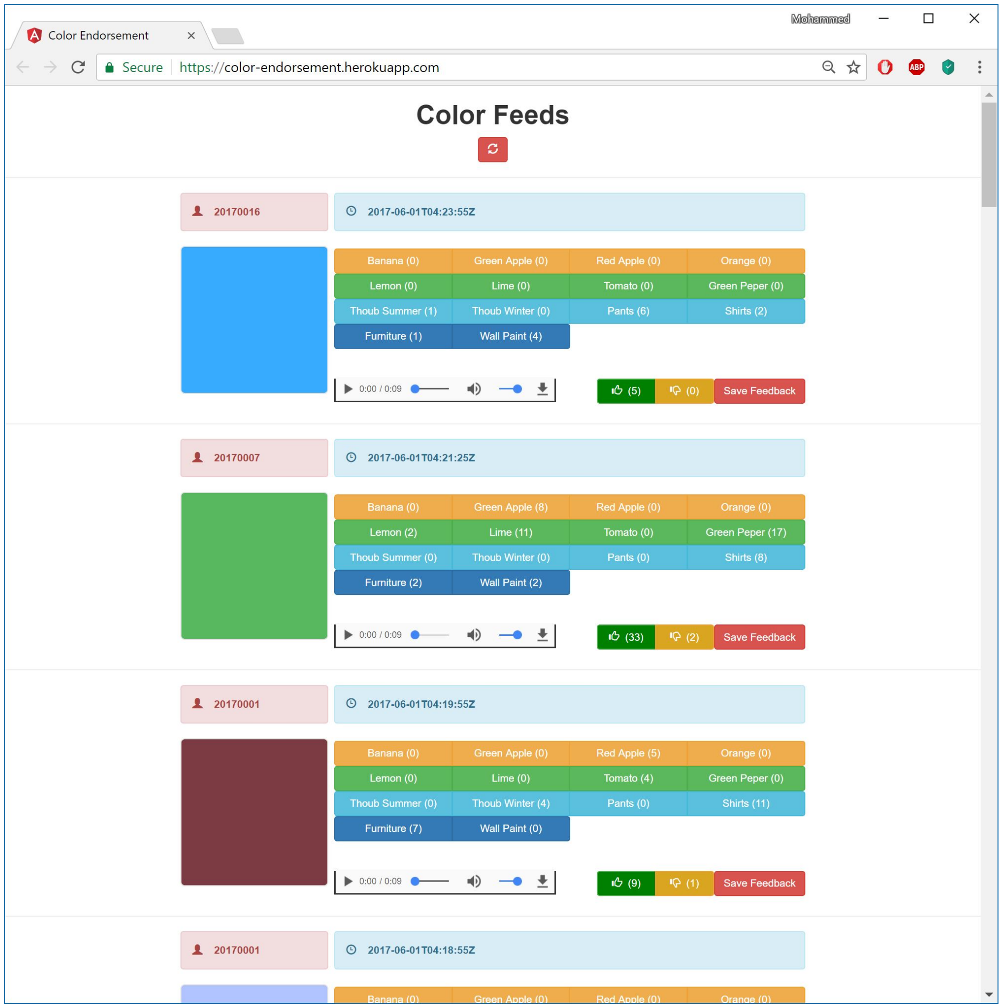

Figure 6.1: Color Endorsement Web Application

comfortable to sighted people. To bridge this possible gap, we developed a web application for color endorsement shown in Figure 6.1.

6.2 Web Application Architecture

As shown in Figure 6.2, this web application consists of client side and server side. The client side is developed by *TypeScript (JavaSript ES6)* using *Angular 4 Framework*, while the server side is developed using *Node.js*.

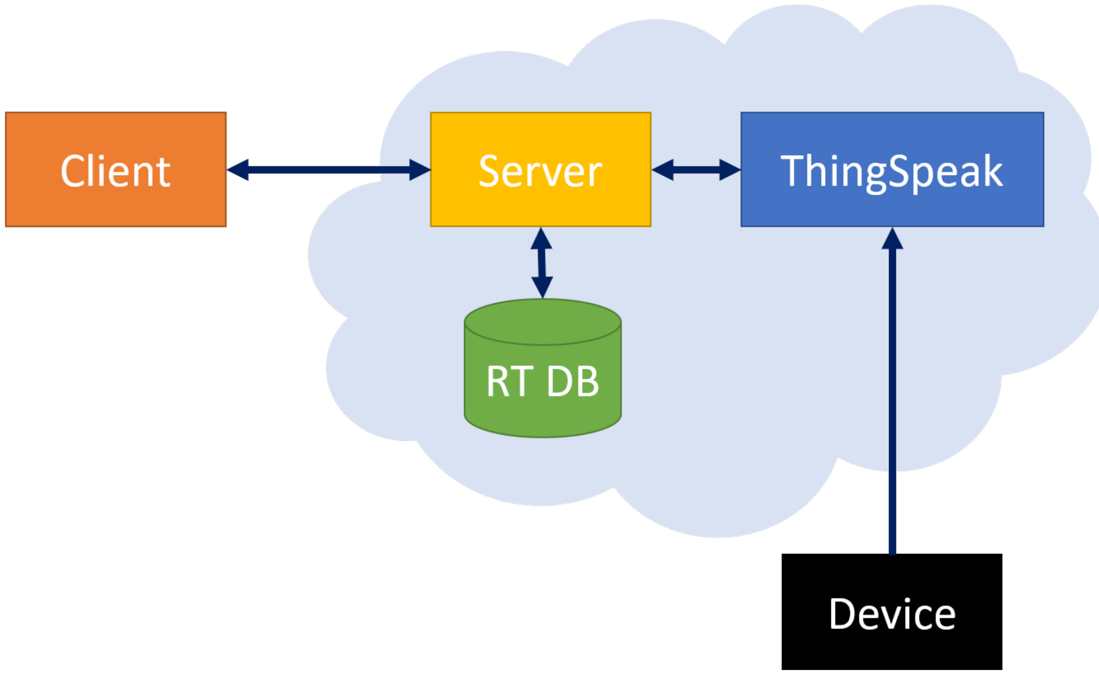

Figure 6.2: Online Service Architecture

The client side runs on the browser and interacts mostly with sighted users. The main page shows the color feeds detected by VI users using our device. Each color feed of four parts: (1) Device ID reporting that color, (2) Time-stamp of the reported color, (3) View of the color, and (4) Color endorsement mechanism. The first three parts are data collected from the server and rendered on the browser. The color endorsement mechanism consists of buttons representing possible uses of the detected color and whether the color is visually comfortable or not. When a sighted user endorses a color, the feedback information is sent to the server.

The server side manages the collected data either from the device or from the client side. The data collected by the server are stored in the Firebase Realtime NoSQL cloud-based database provided by Google. The data are stored and retrieved as JSON format. The device communicates the data of the detected color

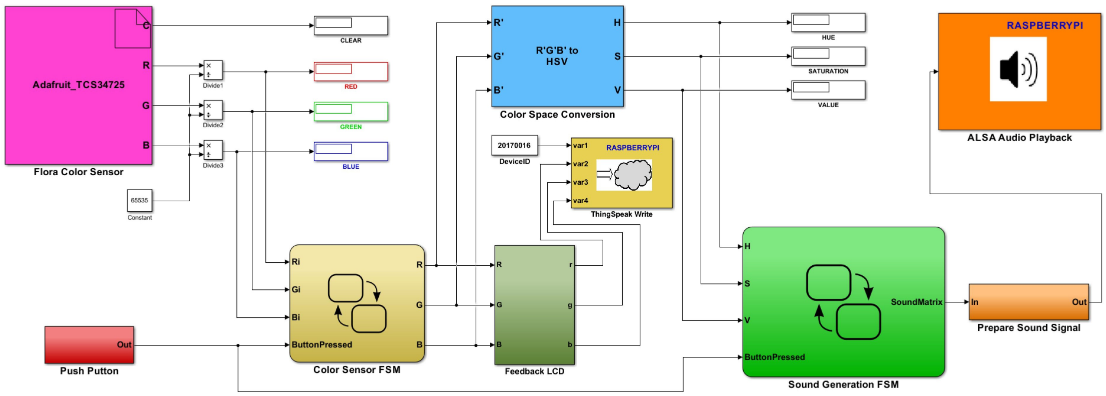

Figure 6.3: Final Matlab Simulink Model

to the server through an open IoT platform called ThingSpeak. This IoT platform is supported by MATLAB Simulink. The server communicates with ThingSpeak cloud through http requests and restful APIs.

The web application is hosted on the *Heroku Cloud Application Platform*. The web application can be accessed on https://color-endorsement.herokuapp.com/.

6.3 Modification to System Design

As shown in Figure 6.3, the system model is upgraded to accommodate the web application by adding the ThingSpeak Write block provided by Simulink. This block enables the system to send related data of the detected color to ThingSpeak platform. The sent data consists of four variables. The first variable contains the device's ID which is hard coded to the device. The other three variables contain the RGB color values of the detected color. Figure 6.4 shows the settings for the ThingSpeak Write block. The write API key is provided by the ThingSpeak platform for registered users among other API keys. The number of variables is

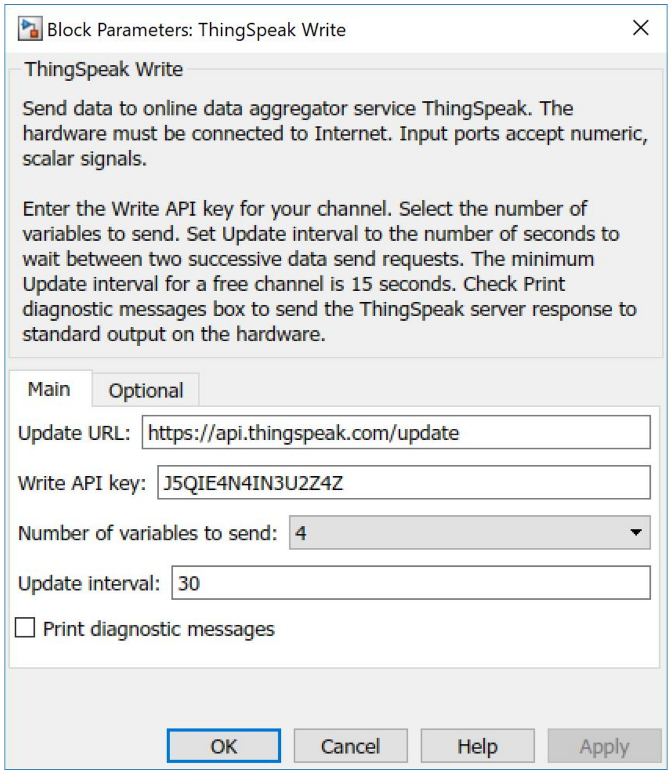

Figure 6.4: ThingSpeak Simulink Block

chosen to be four variables to accommodate the data to be sent. Every 30 seconds, the current data will be sent to ThingSpeak cloud. The 30 second interval is used for the sake of testing.

CHAPTER 7

SYSTEM EVALUATION

In order to assess the effectiveness of our new device, experiments were performed with VI and blindfolded participants. The experiments were carried out over a period of time in a number of separate sessions. The experiments have been performed with four volunteers: a VI volunteer and three blindfolded volunteers. The VI volunteer is an early blind (lost his vision by the age of four) male in his mid-thirties. On the other hand, the blindfolded subjects are males in their late-twenties.

7.1 Introducing the Device Experiment

The following sub-experiments were performed during a three-hour session. The subject was given a short break between the different experiments. Artificial fruits and vegetables were used as shown in figure 7.1.

At the beginning, the artificial fruits and vegetables were introduced to the subject. He was able to successfully recognize the fruits and vegetables using their

Figure 7.1: Artifial fruits for the experiment

shapes and textures with no assistance from our side. The following observations were made about the subject:

- He was able to recognize the orange fruit because of its texture,
- He was not able to recognize the grape because of the way the artificial grape vine is made,
- He was able recognize the green and red apples but he could not recognize their colors. The subject confirmed that he can differentiate between red and green apples through taste because a green apple is more sour,
- He was able to recognize the lemon and tell its color because lemons are yellow while limes are green and a bit smaller,
- He was not able to recognize the eggplant or tell its color because he confirmed that he has never dealt with this kind of fruits in real life.

Figure 7.2: The VI subject performing experiment 1

Next, the device was introduced to the subject and its purpose was explained. Also, the subject was trained on how to correctly use the device, Figure 7.2. The sounds of the colors of the artificial fruits and vegetables used in the experiments were generated and played back via the speaker attached to the device.

7.1.1 Task 1

We started with two color-contrasted objects: (1) Red Tomato and (2) Green Pepper. The subject tried the device with the objects for approximately two minutes. After that, we tested the subject and he could easily recognize the acoustic signal representation for the colors of the objects and even associated the correct color name to them ("red or green").

Figure 7.3: Blindfolded subject performing experiment 1

7.1.2 Task 2

In this experiment, we introduced two objects similar in color: (1) Banana and (2) Lemon. After trying the two objects, the subject could not discriminate between the two colors. He confused the two sounds for the banana. Both are yellow with different intensities.

7.1.3 Task 3

Here, we presented the red and green apples. He immediately recognized the red apple by remembering the sound of red tomato from experiment 1. He was able to recognize and concludes by himself that the apple green is lighter in terms of color than the pepper green by describing the latter as dark green. We were surprised that the subject was able to qualify the green of the pepper as dark green. We thought that he will not be familiar with such terminology because he never used it before.

7.1.4 Task 4

This time, we introduced three objects with three different colors: Banana, Orange, and Red Apple. The subject has been asked to listen to the color sound of each object twice. He was able to identify the orange and the apple (as red) but confused the banana (yellow) for the orange. We think the subject was exhausted at this stage.

7.1.5 Task 5

We provided the subject with the red apple and the banana. Then, we asked him to describe the objects by their shapes and sound. He described the red apple as round with deep sound, while the banana was described as tall with narrow/sharp sound. After that, we gave him an orange and asked him to describe it. His description for the orange was round but not that deep sound in comparison with sound of red in the tomato.

Table 7.1 summarizes the results of the VI volunteer (subject 1) for the first four tasks and also shows the results of the blindfolded volunteer (subjects 2-4) for the same experiments. We can see that the blindfolded subjects took more the training time than the VI subject. In return, the blindfolded subjects had higher success rate than the VI subject. However, the average training time for the volunteers was 11.25 minutes with 90% success rate.

Table 7.1: Experiment 1: Experimental Data

	Volunteers			
	Subject 1	**Subject 2**	**Subject 3**	**Subject 4**
Red Tomato	Success	Success	Success	Success
Green Pepper	Success	Success	Success	Success
Banana	Success	Failure	Success	Failure
Lemon	Failure	Success	Success	Success
Red Apple	Success	Success	Success	Success
Green Apple	Success	Success	Success	Success
Green pepper	Success	Success	Success	Success
Oranges	Success	Success	Success	Success
Banana	Failure	Success	Success	Success
Red Apple	Success	Success	Success	Success
Training Time	6 minutes	12 minutes	15 minutes	12 minutes
Success Rate	80%	90%	100%	90%

7.1.6 Discussion

We noticed that the experiment increased the description language of the VI volunteer for the fruits and vegetables by adding the acoustic description of their colors. Also, he could distinguish between the acoustic representations of similar hues but could not associate them to linguistic names. Hence, We need to enrich their vocabulary to enable them to precisely describe the colors through sounds. In addition, the subject was able to conclude some of the sound parameters and what they represent during the experiments. The third acoustic signal gives indication for the color intensity. Thus, the approach used to represent the colors is intuitive and easy to interpret where it does not require a lot of training.

Based on the feedback received from the VI volunteer, the sound representing the color is pleasant and its loudness is suitable. Finally, the VI volunteer was pleased with the easy-to-use device and enjoyed using it throughout the session

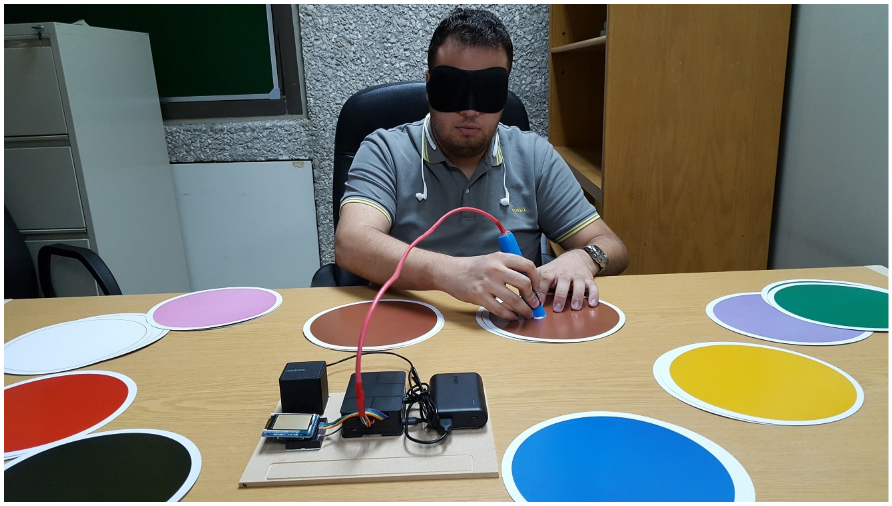

Figure 7.4: Blindfolded subject performing experiment 2

and got interested more in the concept of colors. He even proposed to introduce the device to some of his friends.

7.2 Color Matching Experiment

The objective of this experiment is to test the participants' ability to match similar colors together using the developed device. The experiment follows the following procedure:

1. Two groups of 10 similar colors are introduced to the participants.

2. The participants are asked to match similar colors together and the time needed to complete the task is computed.

Table 7.2: Experiment 2: Experimental Data

Volunteer	**Successful Matches**	**Time**
Subject 2	10	10m 50s
Subject 3	8	15m 30s
Subject 4	10	15m 40s
Average	9.33	14m
Success Rate	**93.33%**	

7.2.1 Discussion

The experiment has been performed by the blindfolded participants. Table 7.2 summarizes the experimental data for three participants. Subject 2 and 4 were able to match up all 10 pairs of cards. Subject 3 missed two pairs of matches. The mismatch were between the yellow card and the orange one. Their acoustic representation were very similar in the hue and value tones but with small difference in the saturation tone. However, the average success rate was 93.33%. It is worth mentioning that no training time was needed to perform this experiment compared to [36] with similar success rate. Also, less time needed to complete the experiment.

7.3 Color Shades Experiment

7.3.1 Color Grouping Task

This task aim is to test the participants' ability to group similar colors with different shades by using the developed device. The task follows the following procedure:

Table 7.3: Task 3a: Experimental Data

Volunteer	**Successful Trials (out of 12)**	**Time**
Subject 2	12	8m 50s
Subject 3	12	11m 30s
Subject 4	11	15m 35s
Average	11.67	11m 40s
Avg. Success Rate	**97.25%**	

Figure 7.5: Blindfolded subject performing color grouping task

1. A group of 12 colored cards are introduced.

2. The participant are asked to group similar color shades. The time needed to complete the task is computed.

7.3.2 Color Shades Ordering Task

The objective of this task is to test the participants' ability to order color shades from a lighter shade to a darker one. The task follows the following procedure:

1. Five color groups consisting of five color shades are introduced.

Table 7.4: Task 3b: Experimental Data

Volunteer	**Successful Trials (out of 18)**	**Time**
Subject 2	18	11m 15s
Subject 3	16	20m 30s
Subject 4	18	15m 20s
Average	17.67	14m
Avg. Success Rate	**96.28%**	

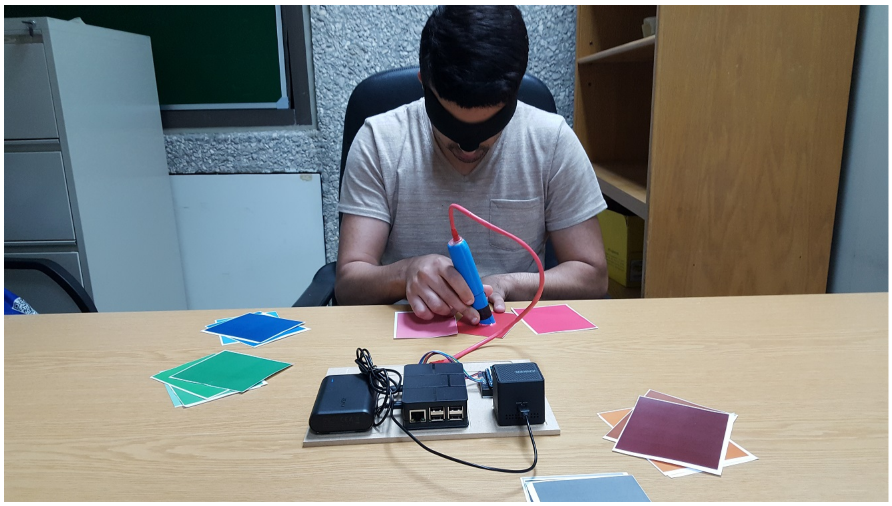

Figure 7.6: Blindfolded subject performing shades ordering task

2. The participant are asked to order color cards according to their shades and the time needed to complete the task is computed.

7.3.3 Discussion

The two tasks of this experiment is performed separately by the blindfolded subjects. As shown in Table 7.3, the first task has been completed in the average of 11 minutes and 40 seconds with success rate of 97.25%. It is worth mentioning that, Subject 2 was able to recognize some of the colors based on his previous experiments. The second task has been performed in average of 14 minutes and

Table 7.5: Experiment 4: Experimental Data

Volunteer	**Training Time**	**Successful Trials**	**Response Time**
Subject 2	12m 45s	39/40	5.14s
Subject 3	15m 10s	40/40	3.66s
Subject 4	5m 25s	40/40	3.27s
Average	11m	39.67/40	4.02s
Success Rate	**99.18%**		

success rate of 96.28% as shown in Table 7.4. Subject 3 missed two very similar green shades, where their acoustic representation were very hard for him to distinguish.

7.4 Color Recognition Experiment

The aim of this experiment is to test the participants' ability to recognize colors using our new color reporting mechanism. This experiment follows the following procedure:

1. The participant is asked to train for 5 colors (red, pink, yellow, green, and blue) and the needed time for training is recorded.

2. Then, the participant is randomly tested for 40 trials from the 5 colors and the response time for every trial is computed.

7.4.1 Discussion

This experiment is performed by the blindfolded subjects. Table 7.5 summarizes the experimental data for the participants. The average training time for this

Table 7.6: Comparison with literature

Year	System Name	Color Space	Color Bandwidth	Training Time	Results
2017	Sonic Color	HSV	All	24m	99.18%
2014 [21]	EyeMusic	HSL	5	120m-180m	81%
2013 [34]	SonarX	HSV	All	10m	53.13%
2012 [22]	-	HSL	All	240m-300m	96%

Table 7.7: Overall Results

Volunteer	Training Time	Successful Trials	Success Rate
Subject 2	24m 45s	88/90	97.78%
Subject 3	30m 10s	86/90	95.56%
Subject 4	17m 25s	88/90	97.78%
Average	24m	87.33/90	97.04%

experiment is 11 minutes with success rate of 99.18%. The average response time of the participants is 4 seconds which is less than the total duration of the acoustic representation of a color (i.e. 6 seconds). Compared to the previous work done in [21] [34] [22], our proposed system on average needs much less training time with higher success rate as shown in Table 7.6.

7.5 Results

By combining the color trials performed by the participants we can get an overall evaluation to our system compared to previous system. Table 7.7 summarizes the experimental data for the four experiments. We can see that, the overall success rate of our proposed system is 97.04% with only 24 minutes of training time. The participants performed total of 90 color trials.

CHAPTER 8

CONCLUSION AND FUTURE WORK

A realtime embedded system to assist visually impaired people recognize colors has been designed and implemented. The proposed system detects colors using a color sensor module. A module is then used to convert the sensed colors from the RGB color space to the HSV color space. the resulting HSV components are then converted through a processing unit into three consecutive audio signals. By this, the first objective of the thesis is achived.

As for the second objective of the thesis, a prototype was built as proof-of-concept for our proposed system using MATLAB Simulink and Raspberry PI 3. The prototype is a pen-like device that has been tested with visually impaired and blindfolded subjects.

To achive the third objective of the thesis, a web application for color endosment was developed and deployed. It helps the VI people integrate more into society by benefiting from the sighted people experience with the detected colors using our system.

A designed set of four experiments have been performed to evaluate our proposed system. The first experiment was to familiarize the participants with proposed system and the implemented device. The average training time for the participants was 11.25 minutes with 90% success rate. The second experiment was for testing the ability of the system in helping the participants in matching similar pairs of colors. The average success rate of the second experiment was 93.33% with no training time needed. The third experiment designed to test the ability of the participants in identifying and recognizing color shades The average success rate for this experiment was 96.77%. The last experiment was to test the ability of the participants to recognize five colors using the proposed system. The average training time for the last experiment was 11 minutes with success rate of 99.18%. By this, the fourth objective of the thesis is met.

The overall success rate of our proposed system was 97.04% with less than 30 minutes of training time. Such outcomes show the robustness and the usefulness of such a system. The device was user friendly and easy to use.

As for the future work, non-linear interpolation for color-to-frequency mapping should be explored and investigated. Also, playback time tuning should be provided to give the VI peopke the option to play generated sounds at faster rates especially after the training stage. In addition, utilizing the forier series with the HSV color model should be investigated. For example, the sqaure, sawtooth or triangle signals can be used to represent Hue, Saturation or Value sounds with the chirping frequency as the fundemental drequency. Furthermore, a mobile application can be developed as an alternative for the device. The mobile application will provide more VI volunteers and wide variety of age groups for the system experimental evaluation.

REFERENCES

[1] J. Dahl and National Eye Institute, “Visual Acuity Testing,” 2008. [Online]. Available: https://www.nei.nih.gov/photo/visual-acuity-testing

[2] J. A. Black and D. S. Hayden, “The Note-Taker: An assistive technology that allows students who are legally blind to take notes in the classroom,” *Computer Vision and Pattern Recognition Workshops CVPRW 2010 IEEE Computer Society Conference on*, no. 1, pp. 131–137, 2010. [Online]. Available: http://portal.acm.org/citation.cfm?doid=1878803.1878828

[3] J. Jungil, Y. Hongchan, L. Hyelim, and C. Jinsoo, “Graphic haptic electronic board-based education assistive technology system for blind people,” *2015 IEEE International Conference on Consumer Electronics (ICCE)*, pp. 364–365, 2015.

[4] S. Horvath, J. Galeotti, B. Wu, R. Klatzky, M. E. L. S. Fellow, and G. S. Member, “FingerSight: Fingertip Haptic Sensing of the Visual Environment,” *IEEE Journal of Translational Engineering in Health and Medicine*, vol. 2, no. February, 2014.

[5] “Colorblindness and Achromatopsia,” 2013. [Online]. Available: http://www.achromatopsia.info/color-blindness

[6] “Color Teller.” [Online]. Available: http://www.brytech.com/colorteller/index.htm

[7] “Talking Thermostats.” [Online]. Available: https://www.talkingthermostats.com/blind.shtml

[8] “iBill - Talking Banknote Identifier.” [Online]. Available: http://www.orbitresearch.com/ibill_details.php

[9] Centers for Disease Control and Prevention, “Blindness and Vision Impairment,” 2017. [Online]. Available: https://www.cdc.gov/healthcommunication/ToolsTemplates/EntertainmentEd/Tips/Blindness.html

[10] J. Marsden, S. Stevens, and A. Ebri, “How to Measure Distance Visual Acuity,” *Community Eye Health 27.85*, vol. 27.85, 2014.

[11] American Optometric Association, “Visual Acuity: What is 20/20 Vision?” 2010. [Online]. Available: https://www.aoa.org/patients-and-public/eye-and-vision-problems/glossary-of-eye-and-vision-conditions/visual-acuity

[12] World Health Organization, “Vision impairment and blindness,” 2017. [Online]. Available: http://www.who.int/mediacentre/factsheets/fs282/en

[13] ——, "10 facts about blindness and visual impairment," 2014. [Online]. Available: http://www.who.int/features/factfiles/blindness/en/

[14] M. A.Hersh and M. A.Johnson, *Assistive Technology for Visually Impaired and Blind People*, 2008.

[15] K. Vogt, "a Quantitative Evaluation Approach To Sonifications," in *The 17th International Conference on Auditory Display (ICAD-2011)*, Budapest, Hungary, 2011.

[16] E. R. Johnson, "Sensation and perception," 2016.

[17] S. Aukstakalnis, *Practical Augmented Reality: A Guide to the Technologies, Applications, and Human Factors for AR and VR*, ser. Usability. Pearson Education, 2016. [Online]. Available: https://books.google.com.sa/books?id=4oz_DAAAQBAJ

[18] S. Bhatlawande, M. Mahadevappa, J. Mukherjee, M. Biswas, D. Das, and S. Gupta, "Design, development, and clinical evaluation of the electronic mobility cane for vision rehabilitation," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 2014.

[19] S. Mascetti, L. Picinali, A. Gerino, D. Ahmetovic, and C. Bernareggi, "Sonification of guidance data during road crossing for people with visual impairments or blindness," *International Journal of Human Computer Studies*, vol. 85, pp. 16–26, 2016. [Online]. Available: http://dx.doi.org/10.1016/j.ijhcs.2015.08.003

[20] N. Z. Salamon, J. C. S. J. Junior, and S. R. Musse, "Seeing the Movement through Sound : Giving Trajectory Information to Visually Impaired People," in *Brazilian Symposium on Games and Digital Entertainment, SBGAMES*, vol. 2014-Decem, no. December, 2014, pp. 12–17.

[21] S. Abboud, S. Hanassy, S. Levy-Tzedek, S. Maidenbaum, and A. Amedi, "EyeMusic: Introducing a 'visual' colorful experience for the blind using auditory sensory substitution," *Restorative Neurology and Neuroscience*, vol. 32, no. 2, pp. 247–257, 2014.

[22] M. Banf and V. Blanz, "A modular computer vision sonification model for the visually impaired," in *Proceedings of the 18th International Conference on Auditory Display (ICAD 2012)*, Atlanta, GA, USA, 2012, pp. 121–128.

[23] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, third edition ed. Pearson Education, Inc., 2008.

[24] E. A. Lee and S. A. Seshia, *Introduction to Embedded Systems - A Cyber-Physical Systems Approach*, second edition ed. MIT Press, 2017.

[25] G. P. Garcia, J. O. Bejar, and F. M. Miramontes, "A Prototype Helping Device for the Visually Impaired Using an Optical to Mechanical Transducer," 2013.

[26] D. A. Cazan, R. Vârbanescu, and D. Popescu, "Algorithms and Techniques for Image to Sound Conversion for Helping the Visually Impaired People Application Proposal."

[27] E. Peng, P. Peursum, and L. Li, "Product Barcode and Expiry Date Detection for the Visually Impaired Using a Smartphone," *2012 International Conference on Digital Image Computing Techniques and Applications (DICTA)*, pp. 1–7, 12 2012. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6411673

[28] B. Taylor and D.-j. Lee, "Smart Phone-based Indoor Guidance System for the Visually Impaired," vol. 2012, no. December, pp. 5–7, 2012.

[29] L. Balduzzi, G. Fusco, F. Odone, S. Dini, M. Mesiti, a. Destrero, and a. Lovato, "Low-cost face biometry for visually impaired users," *2010 IEEE Workshop on Biometric Measurements and Systems for Security and Medical Applications*, pp. 45–52, 9 2010. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5610444

[30] P. E. Lanigan, A. M. Paulos, A. W. Williams, and P. Narasimhan, "Trinetra : Assistive Technologies for the Blind," *CyLab*, pp. 1–2, 2006.

[31] L. David, R. Vasconcelos, L. Alves, R. Andre, G. Baptista, and M. Endler, "A Communication Middleware for Scalable Real-Time Mobile Collaboration," *2012 IEEE 21st International Workshop on Enabling Technologies: Infrastructure for Collaborative Enterprises*, pp. 54–59, 6 2012. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6269699

[32] D. T. V. Pawluk, R. J. Adams, S. Member, and R. Kitada, "Designing Haptic Assistive Technology for Individuals Who Are Blind or Visually Impaired," vol. 8, no. 3, pp. 258–278, 2015.

[33] S. Cavaco, J. T. Henriques, M. Mengucci, N. Correia, and F. Medeiros, "Color Sonification for the Visually Impaired," in *Procedia Technology*, vol. 9. Elsevier B.V., 2013, pp. 1048–1057. [Online]. Available: http://dx.doi.org/10.1016/j.protcy.2013.12.117

[34] S. Cavaco, M. Mengucci, J. T. Henriques, N. Correia, and F. Medeiros, "From pixels to pitches: Unveiling the world of color for the blind," in *SeGAH 2013 - IEEE 2nd International Conference on Serious Games and Applications for Health, Book of Proceedings*, 2013.

[35] M. Banf and V. Blanz, "Sonification of images for the visually impaired using a multi-level approach," in *Proceedings of the 4th Augmented Human International Conference on - AH '13*, 2013, pp. 162–169.

[36] G. Bologna, B. Deville, M. Vinckenbosch, and T. Pun, "A Perceptual Interface for Vision Substitution in a Color Matching Experiment," in *2008 International Joint Conference on Neural Networks (IJCNN 2008)*. Hong Kong: IEEE, 2008, pp. 1622–1629.

[37] "Al-Watan Online," 2013. [Online]. Available: http://www.alwatan.com.sa/Culture/News_Detail.aspx?ArticleID=159963\&CategoryID=7

[38] K. v. d. Doel, "SOUNDVIEW: SENSING COLOR IMAGES BY KINESTHETIC AUDIO," in *Proceedings of the 2003 International Conference on Auditory Display (ICAD03)*, Boston, MA, USA, 2003, pp. 303–306.

[39] D. S. Burch, R. Va, and D. T. V. Pawluk, "A Cheap , Portable Haptic Device for a Method to Relay 2-D Texture-Enriched Graphical Information to Individuals who are Visually Impaired," in *ASSETS09*, Pittsburgh, Pennsylvania, USA, 2009, pp. 215–216.

[40] T. Wright and J. Ward, "The evolution of a visual-to-auditory sensory substitution device using interactive genetic algorithms," *The Quarterly Journal of Experimental Psychology*, vol. 66, no. 8, pp. 1620–1638, 2013.

[41] Dr. Ir. Stéphane Pigeon, "Low Frequency Range Test." [Online]. Available: https://www.audiocheck.net/audiotests_frequencychecklow.php

Vitae

- **Name:** Mohammed Tawfiq Samara
- **Nationality:** Palestinian
- **Date of Birth:** 2^{nd} of August, 1987
- **Email:** *mtsamara@gmail.com*
- **Permenant Address:** 144/12 Shujaiya, Gaza Strip, Palestine
- **Current Address:** KFUPM, Dhahran, Saudi Arabia

- **Education:**

  1. **Islamic University - Gaza (Gaza Strip, Palestine)**

     B.Sc. in Computer Engineering - GPA 89.14% (June 2011)

  2. **Abu Arish Secondary School (Jazan, Saudi Arabia)**

     General Secondary Certificate - GPA 99.25% (May 2005)

- **Conferences:**

  1. **4^{th} IEEE International Conference on Engineering Technologies and Applied Sciences (ICETAS 2017) (Bahrain) (November 2017)**

Mohammed Samara, Jihad AlSadah, Mohamed Deriche, and Yahya Osais, "A Color Recognition System for the Visually Impaired People" *(Best Paper Award)*

2. **5^{th} IUG Engineering Day (Gaza Strip, Palestine) (May 2010)**

   Abdelghani Abujabal, Elias Fashho, and Mohammed Samara, "Microcontroller Based Taximeter"

- **Projects:**

  1. **Color Endorsement (Web Application)**

     Design, and Development of the web application using Angular and Node.js (2017)

  2. **Color Recognition for the Blind (Raspberry Pi-Based Device)**

     Building a device that is using a Raspberry Pi, a color sensor, a graphical LCD, and a speaker (2017)

  3. **Save File to Dropbox (Website)**

     Design, and Development of the website using Node.js and Bootstrap (2015)

  4. **Microcontroller based Taximeter (Embedded System)**

     Analysis, Design, and Development of the system using MIKROBASIC and PIC MCU (2011)